

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2003, 17:14:42 : Search time 87 Seconds
(without alignments)
490.775 Million cell updates/sec

Title: US-09-852-100A-2

Perfect score: 1439

Sequence: 1 MHILKSPNVPRAHGQKNT.....TLRLSLITNEIRKTLQYP 269

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1439	100.0	269	AAW94291	Human beta-amyloid
2	1439	100.0	269	AAV70759	Human beta-amyloid
3	1439	100.0	269	AAE33877	Human BBP-1 protei
4	748	52.0	139	AAV12358	Human 5' EST secre
5	673.5	46.8	162	AAV36021	Extended human sec
6	667.5	46.4	148	AAV12426	Human 5' EST secre
7	338	23.5	178	ABB65236	Drosophila melanog
8	293	20.4	100	AAU97631	RNA polymerase II
9	201	14.0	221	ABR47818	Human secreted pro

10	201	14.0	221	24	ABR00112	Human gene 102 enc
11	201	14.0	222	20	AAV27658	Human secreted pro
12	200	13.9	221	21	AAV70761	Human beta-amyloid
13	200	13.9	221	24	ABR48493	Human Amyloid Apop
14	200	13.9	247	21	AAE08641	A human TANGO 223
15	200	13.9	247	21	AAE08656	A human TANGO 223
16	200	13.9	247	21	AAE08657	A human TANGO 223
17	200	13.9	247	21	AAE08658	A human TANGO 223
18	196	13.6	230	21	AAE08645	DNA encoding a mur
19	182	12.6	284	22	ABE59014	Drosophila melanog
20	152.5	10.6	225	22	ABE11574	Human secreted pro
21	138.5	9.6	214	21	AAE87351	Human signal pepti
22	136.5	9.5	214	21	AAV70760	Human beta-amyloid
23	136.5	9.5	224	21	AAE43573	Human cancer assoc
24	132	9.2	224	22	ABE65466	Drosophila melanog
25	121	8.4	171	22	AAE67818	Human guanosine li
26	121	8.4	171	22	AAE93475	Human polypeptide,
27	121	8.4	171	22	AAE93578	Human polypeptide,
28	121	8.4	171	22	AAE81266	Human AFP protei
29	121	8.4	171	23	ABE90376	Human polypeptide
30	121	8.4	172	20	AAE45270	Human secreted pro
31	119	8.3	140	23	ABE40585	Human secreted pro
32	115.5	8.0	213	23	ABE72307	Rat protein isolat
33	95.5	6.6	944	22	AAE32044	Human vesicle-assc
34	90.5	6.3	341	22	AAE41757	Human polypeptide
35	88.5	6.2	224	22	ABE20306	Novel human diagn
36	88	6.1	68	24	AAE33878	Human BBP-1 protei
37	88	6.1	764	20	AAE94360	Human Factor B and
38	86.5	6.0	316	22	ABE64318	Drosophila melanoc
39	86	6.0	764	20	AAE94365	Human Factor B and
40	86	6.0	764	20	AAE94372	Human Factor B and
41	85	5.9	310	17	AAE05156	Rat mature LAMP.
42	85	5.9	338	17	AAE05153	Rat LAMP residues
43	85	5.9	338	17	AAE05154	Rat LAMP residues
44	85	5.9	576	21	AAE75498	Neisseria gonorrhoe
45	85	5.9	576	24	ABE80808	N. gonorrhoeae am;

ALIGNMENTS

RESULT 1
AAW94291
ID AAW94291 standard; Protein; 269 AA.
XX AAW94291;
AC AAW94291;
XX 27-APR-1999 (first entry)
XX Human beta-amyloid peptide-binding protein (BBP).
DE Human beta-amyloid peptide-binding protein (BBP).
XX Beta-amyloid peptide binding protein; BBP; beta-amyloid protein; BAP;
KW human; Alzheimer's disease.
XX Homo sapiens.
XX Key Location/Qualifiers
XX Region 68..269
XX FT /note- "specifically claimed fragment having
XX FT beta-amyloid peptide binding activity."
XX
XX W05846636-A2.
XX
XX PD 22-OCT-1998.
XX
XX PF 14-APR-1998; 98WC-US07462. ✓
XX
XX PR 16-APR-1997; 97US-0064583.
XX
XX PA (AMHP) AMERICAN HOME PROD CORP.
XX
XX V1 Bard JA, Jacobsen JS, Kajkowski EM, Ozenberger BA;
XX Walker SG;

```
XX WPI: 1999-080736/G7.
DR N-PSDB: AAX05735.
XX
XX Polynucleotide encoding beta-amyloid peptide binding protein - used
PT to identify inhibitors of beta-amyloid peptide for treating
PT Alzheimer's disease
XX
XX Claim 7: Pages 43-44; 59pp; English.
XX
XX The present sequence represents a beta-amyloid peptide binding protein
CC (BBP). The polynucleotide comprising the entire BBP nucleotide sequence
CC of clone BBPI-f1 is deposited under the accession number ATCC 98617. The
CC polynucleotide comprising a fragment of BBP (nucleotides 202-807 of the
CC full length BBP) of clone PEK196 is deposited as ATCC 98395. Host cells
CC transformed with a vector comprising the BBP nucleic acid are used for
CC the recombinant production of the protein. The protein can be used in a
CC method for diagnosing a disease characterised by aberrant expression of
CC human beta-amyloid protein (BAP). The protein can also be used in a
CC method for screening for compounds which regulate expression of a BAP
CC binding protein. The proteins, antibodies and identified compounds can be
CC used in the treatment or prevention of Alzheimer's disease.
XX
XX Sequence 269 AA:
XX
XX Query Match 100.0%; Score 1439; DB 20; Length 269;
XX Best Local Similarity 100.0%; Pred. No. 9.4e-142;
XX Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MHILKGSPNVIPRAHGQKNTRRDGTGLYPMRGPFKNLALLPFLGSGSGSGEKVSV 60
DB 1 MHILKGSPNVIPRAHGQKNTRRDGTGLYPMRGPFKNLALLPFLGSGSGSGEKVSV 60
QY 61 SKMAAAMPSPGSAPEAVTARLVGLWLFVSVTTGPGAVATSGAGEESLKCEDLKVGQYIC 120
DB 61 SKMAAAMPSPGSAPEAVTARLVGLWLFVSVTTGPGAVATSGAGEESLKCEDLKVGQYIC 120
QY 121 KDPKINDATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFPPKPSICRNNG 180
DB 121 KDPKINDATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFPPKPSICRNNG 180
QY 181 YSYKVAVALSLFLGWLGDADRFYLGYPALGLLKCTVGFCGIGSLIDFLILSMQIVGPSDG 240
DB 181 YSYKVAVALSLFLGWLGDADRFYLGYPALGLLKCTVGFCGIGSLIDFLILSMQIVGPSDG 240
QY 241 SSVIIDYGTGTRLTSLTNETFRKTQLYP 269
DB 241 SSVIIDYGTGTRLTSLTNETFRKTQLYP 269
XX
XX RESULT 2
XX ID AAY70759 standard: Protein; 269 AA.
XX AC AAY70759;
XX
XX 24-JUL-2000 (first entry)
XX
XX Human beta-amyloid peptide (BAP) binding protein, BBP1.
XX
XX Beta-amyloid peptide binding protein; BBP; BAP; tumour; suppressor;
KW G-protein coupled receptor; GPCR; integral membrane protein; antigen;
KW neuronal cell; nonhuman primate; NHP; G-protein signalling pathway;
KW apoptosis; immunogen; therapeutic; treatment; prevention; diagnostic.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Domain 177..198
XX /label= Transmembrane_domain_1
XX Domain 199..201
XX /label= DRF_motif
XX /note= "Substitution of the Arg abrogates protection"
```

```
FT Domain 213..238
FT /label= Transmembrane_domain_2
XX
XX WQ20002125-A2.
XX
XX 20-APR-2000.
XX
XX 13-OCT-1999; 99WC-US21521.
XX
XX 13-OCT-1998; 98US-0104104.
XX
XX (AMHP ) AMERICAN HOME PROD CORP.
XX
XX ozenberger BA, kajkowski EM, LO CF;
XX
XX WPI: 2000-317982/27.
XX N-PSDB: AAE52369.
XX
XX Novel G-protein-coupled receptor-like proteins and polynucleotides
XX useful for regulating apoptosis, comprises integral membrane protein
XX traversing the membrane twice.
XX
XX Example 1: Page 62-63; 68pp; English.
XX
XX The present sequence is the beta-amyloid peptide (BAP) binding protein-1
XX (BBP1). It is an integral membrane protein, that traverse the membrane
XX twice. It is related to G protein-coupled receptor (GPCR) protein
XX superfamily. It interacts with G-alpha proteins and regulates the
XX activity of G-protein signalling pathways. BBP genes are widely expressed
XX in neuronal cells of nonhuman primate (NHP) brain and overexpressed in
XX some tumours. It functions as a suppressor of apoptosis induction. BBP
XX proteins are used as immunogens to raise antibodies, useful as
XX therapeutics and as antigens in solid phase assays. They are also useful
XX as reagents to identify molecules which effect the interaction of BBP and
XX a cloned protein, that are useful in the treatment or prevention of
XX diseases associated with apoptosis. The polynucleotides are useful for
XX diagnostics.
XX
XX Note: In claim 5, the patent claims an amino acid sequence from figure 2.
XX However, figure 2 does not contain any sequence. It is inferred from the
XX disclosure that the figure 2 sequence refers to BBP1 protein, shown in
XX this sequence.
XX
XX Sequence 269 AA:
XX
XX Query Match 100.0%; Score 1439; DB 21; Length 269;
XX Best Local Similarity 100.0%; Pred. No. 9.4e-142;
XX Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MHILKGSPNVIPRAHGQKNTRRDGTGLYPMRGPFKNLALLPFLGSGSGSGEKVSV 60
DB 1 MHILKGSPNVIPRAHGQKNTRRDGTGLYPMRGPFKNLALLPFLGSGSGSGEKVSV 60
QY 61 SKMAAAMPSPGSAPEAVTARLVGLWLFVSVTTGPGAVATSGAGEESLKCEDLKVGQYIC 120
DB 61 SKMAAAMPSPGSAPEAVTARLVGLWLFVSVTTGPGAVATSGAGEESLKCEDLKVGQYIC 120
QY 121 KDPKINDATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFPPKPSICRNNG 180
DB 121 KDPKINDATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFPPKPSICRNNG 180
QY 181 YSYKVAVALSLFLGWLGDADRFYLGYPALGLLKCTVGFCGIGSLIDFLILSMQIVGPSDG 240
DB 181 YSYKVAVALSLFLGWLGDADRFYLGYPALGLLKCTVGFCGIGSLIDFLILSMQIVGPSDG 240
QY 241 SSVIIDYGTGTRLTSLTNETFRKTQLYP 269
DB 241 SSVIIDYGTGTRLTSLTNETFRKTQLYP 269
XX
XX RESULT 3
XX ID AAE33877 standard: Protein; 269 AA.
XX
```

```

AC AAE33877;
XX
DT 02-MAY-2003 (first entry)
XX
DE Human BBP-1 protein.
XX
KW Human; beta-amyloid peptide-binding protein; BAP; Abeta; betaA2P; BBP;
KW Alzheimer's disease; AD; transgenic; transgenic animal; gene therapy;
KW neuroprotective; neurotropic.
XX
OS Homo sapiens.
XX
PN W0200290499-A2.
XX
PD 14-NOV-2002.
XX
PF 06-MAY-2002; 2002WO-US14223.
XX
PR 09-MAY-2001; 2001US-0852100.
XX
PA (AMBP ) WYETH.
XX
PI Ozenberger BA, Bard JA, Kajkowski EM, Jacobsen JS, Walker SG;
PI Sofia HJ, Howland DS;
XX
DR WPI; 2003-120537/11.
DR N-PSDB; AAD51940.
XX
PT New human beta-amyloid peptide-binding protein, useful for diagnosing
PT and/or treating diseases associated with aberrant expression of
PT beta-amyloid peptide, e.g. Alzheimer's disease.
XX
PS Claim 4; Page 84-85; 85pp; English.
XX
CC The present invention relates to novel human beta-amyloid peptide (BAP;
CC Abeta, betaA2P)-binding (BBP) proteins and polynucleotides encoding such
CC proteins. BBP sequences are useful to diagnose and/or treat diseases
CC associated with aberrant expression of human BAP such as Alzheimer's
CC disease (AD). They are used to generate transgenic animals. Sequences
CC of the invention are also used in gene therapy. The present sequence
CC is human BAP-1 protein.
XX
SQ Sequence 269 AA;
Query Match 100.0%; Score 1439; DB 24; Length 269;
Best Local Similarity 100.0%; Pred. No. 9.4e-142;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MHILKSPNVIPIRAHQKNIIRDGTCLYPMRGPPKKNLALLPFLPLGGSGSGGKYSV 60
Dd 1 MHILKSPNVIPIRAHQKNIIRDGTCLYPMRGPPKKNLALLPFLPLGGSGSGGKYSV 60
Qy 61 SKMAAARSPGSAPEAVTARLVGVLMFVSVTTPGPGAVATAGGSESLKCEDLKVGVIC 120
Dd 61 SKMAAARSPGSAPEAVTARLVGVLMFVSVTTPGPGAVATAGGSESLKCEDLKVGVIC 120
Qy 121 KDPKINDATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFEPKISCRNVNG 180
Dd 121 KDPKINDATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFEPKISCRNVNG 180
Qy 181 YSYKVAVALSLFLGLGADRFYLVGYPALGLLKFCTVPGFCG;GSLIDFIIISMQIVGPSDG 240
Dd 181 YSYKVAVALSLFLGLGADRFYLVGYPALGLLKFCTVPGFCG;GSLIDFIIISMQIVGPSDG 240
Qy 241 SSIYIDYVGTGTRLTSLISINETPRKTLQYP 269
Dd 241 SSIYIDYVGTGTRLTSLISINETPRKTLQYP 269

```

```

RESULT 4
AAV12358
ID AAY12358 standard; Protein: 139 AA.
XX

```

```

AC AAY12358;
XX
DT 17-JUN-1999 (first entry)
XX
DE Human 5' EST secreted protein SEQ ID NO:389.
XX
KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition.
XX
OS Homo sapiens.
XX
PR W09906548-A2.
XX
PD 11-FEB-1999.
XX
PF 31-JUL-1998; 98WO-IB01222.
XX
PR 01-AUG-1997; 97US-0905135.
XX
PA (GEST ) GENSET.
XX
PI Ducicert A, Dumas Milne Edwards J, Lacroix B;
XX
DR WPI; 1999-153778/13.
DR N-PSDB; AAX41191.
XX
PT New nucleic acids encoding human secreted proteins - obtained from
PT cDNA libraries prepared from e.g. liver, ovary, brain, prostate,
PT kidney, lung, umbilical cord, placenta and colon tissue
XX
PS Claim 27; Page 714-715; 824pp; English.
XX
AA AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for
XX human secreted proteins, and encode the proteins given in AAY12261 to
XX AAY12514, respectively. The proteins given represent the signal peptide
XX and an N-terminal fragment of a secreted protein. The nucleic acid
XX sequences can be used for producing secreted human gene products. They
XX can also be used to develop products for diagnosis and therapy. The
XX proteins obtained may have cytokine activity, cell
XX proliferation/differentiation activity, haematopoiesis regulating
XX activity, tissue growth regulating activity, reproductive hormone
XX regulating activity, chemotactic/chemokinetic activity, haemostatic and
XX thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX activity, tumour inhibition activity or other activities. The products
XX can be used in forensic, gene therapy and chromosome mapping procedures.
XX The sequences can also be used for obtaining corresponding promoter
XX sequences. The nucleic acids encoding the signal peptide can be used for
XX directing extracellular secretion of a polypeptide or the insertion of a
XX polypeptide into a membrane, or importing a polypeptide into a cell.
XX
SQ Sequence 139 AA;
Query Match 52.0%; Score 748; DB 20; Length 139;
Best Local Similarity 99.3%; Pred. No. 6.2e-70;
Matches 138; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 63 MAAAWSPGSAPEAVTARLVGVLMFVSVTTPGPGAVATAGGSESLKCEDLKVGVIC 120
Dd 1 MAAAWSPGSAPEAVTARLVGVLMFVSVTTPGPGAVATAGGSESLKCEDLKVGVIC 60
Qy 123 PKINDATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFEPKISCRNVNG 180
Dd 61 PKINDATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFEPKISCRNVNG 120
Qy 183 YKVAVALSLFLGLGADRF 201
Dd 121 YKVAVALSLFLGLGADRF 139

```

RESULT 5

AA136021
ID AAY36021 standard; Protein: 162 AA.
XX AC AAY36021;
XX DI 13-SEP-1999 (first entry)
XX DE Extended human secreted protein sequence, SEQ ID NO. 406.
XX KW Secreted protein; human; cytokine; cellular proliferation; cell movement;
KW cellular differentiation; immune system regulator; anti-inflammatory;
KW hematopoiesis regulator; tissue growth regulator; tumour inhibitor;
KW reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;
KW genetic disease.
XX OS Homo sapiens.
XX PN W09931236-A2.
XX PD 24-JUN-1999.
XX PF 17-DEC-1998; 98WO-IB02122.
XX PR 10-AUG-1998; 98US-0096116.
XX PR 17-DEC-1997; 97US-0069957.
XX PR 09-FEB-1998; 98US-0074121.
XX PR 13-APR-1998; 98US-0081563.
XX PA (GEST) GENSET.
XX PI Bougueleret L, Duclert A, Dumas Milne Edwards J;
XX DR WPI: 1999-385906/32.
XX DR N-PSDB; AAX97705.
XX PT New isolated human secreted proteins
XX PS Claim 9; Page 346-347; 516pp; English.
XX CC This sequence is encoded by an extended human secreted protein coding
CC sequence of the invention. The secreted proteins can be used in treating
CC or controlling a variety of human conditions. The secreted proteins may
CC act as cytokines or may affect cellular proliferation or differentiation.
CC or may act as immune system regulators, hematopoiesis regulators, tissue
CC growth regulators, regulators of reproductive hormones or cell movement
CC or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or
CC tumour inhibition activity. The DNAs can be used in forensic procedures
CC to identify individuals or in diagnostic procedures to identify
CC individuals having genetic diseases resulting from abnormal expression of
CC the genes corresponding to the extended cDNAs. They are also useful for
CC constructing a high resolution map of the human chromosomes. They can
CC also be used for gene therapy to control or treat genetic diseases.
XX SQ Sequence 162 AA;

Query Match 46.8%; Score 673.5; DB 20; Length 162;
Best Local Similarity 84.1%; Pred. No. 4.8e-62;
Matches 127; Conservative 4; Mismatches 17; Indels 3; Gaps 2;
QY 63 MAAWSPGSAPEAVTARLVGLVFWVSVTTPGWAVATSGAESLKCEDLKVGQYICKD 122
DB 1 MAAWSPGSAPEAVTARLVGLVFWVSVTTPGWAVATSGAESLKCEDLKVGQYICKD 60
QY 123 PKINDATQEPVNCNTNYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFKPISCRNVNGYS 182
DB 61 PKINDATQEPVNCNTNYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFKPISCRNVNGYS 120
QY 183 YKVAVALSLFLGLMGADRFLYLG- PALGLL 212
DB 121 YNEQSHVS--FSNMVGSRSILPWPCFGFKV 149

RESULT 6

AA12426
ID AAY12426 standard; Protein: 148 AA.
XX AC AAY12426;
XX DI 17-JUN-1999 (first entry)
XX DE Human 5' EST secreted protein SEQ ID NO:457.
XX KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; hematopoiesis regulator; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition.
XX OS Homo sapiens.
XX PN W09906548-A2.
XX PD 11-FEB-1999.
XX PF 31-JUL-1998; 98WO-IB01222.
XX PR 31-AUG-1997; 97US-0905135.
XX PA (GEST) GENSET.
XX PI Duclert A, Dumas Milne Edwards J, Lacroix B;
XX DR WPI: 1999-153778/13.
XX DR N-PSDB; AAX41259.
XX PT New nucleic acids encoding human secreted proteins - obtained from
PT cDNA libraries prepared from e.g. liver, ovary, brain, prostate,
PT kidney, lung, umbilical cord, placenta and colon tissue
XX PS Claim 27; Page 763-764; 824pp; English.
XX CC AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for
CC human secreted proteins, and encode the proteins given in AAY12261 to
CC AAY12514, respectively. The proteins given represent the signal peptide
CC and an N-terminal fragment of a secreted protein. The nucleic acid
CC sequences can be used for producing secreted human gene products. They
CC can also be used to develop products for diagnosis and therapy. The
CC proteins obtained may have cytokine activity, cell
CC proliferation/differentiation activity, hematopoiesis regulating
CC activity, tissue growth regulating activity, reproductive hormone
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptide can be used for
CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell.
XX SQ Sequence 148 AA;

Query Match 46.4%; Score 667.5; DB 20; Length 148;
Best Local Similarity 83.3%; Pred. No. 1.8e-61;
Matches 125; Conservative 4; Mismatches 18; Indels 3; Gaps 2;
QY 63 MAAWSPGSAPEAVTARLVGLVFWVSVTTPGWAVATSGAESLKCEDLKVGQYICKD 122
DB 1 MAAWSPGSAPEAVTARLVGLVFWVSVTTPGWAVATSGAESLKCEDLKVGQYICKD 60
QY 123 PKINDATQEPVNCNTNYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFKPISCRNVNGYS 182
DB 61 PKINDATQEPVNCNTNYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFKPISCRNVNGYS 120
QY 183 YKVAVALSLFLGLMGADRFLYLG- PALGLL 212

Db 121 YXXQXVS--FSMWGSRSLPWPCEGFV 148

RESULT 7

ABR65236
ID ABR65236 standard; Protein: 178 AA.

XX AC ABR65236;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 22500.

XX KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.

XX PR 23-MAR-2000; 2000US-191637P.

XX PR 11-JUL-2000; 2000US-0614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI: 2001-656860/75.

XX DR N-PSDB: ABL09339.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.

XX PS Disclosure; SEQ ID NO 22500; 2lpp + Sequence Listing; English.

CC CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA sequences (ABU1840-ABU16175) and the encoded proteins (ABB57737-ABB72072).

CC CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 178 AA;

Query Match 23.5%; Score 338; DB 22; Length 178;

Best Local Similarity 42.6%; Pred. No. 6.5e-27;

Matches 69; Conservative 30; Mismatches 49; Indels 14; Gaps 5;

QY 107 SLKCEDLK-VGOYICKDP---KINDATQEPVNCIN--TAHVCGFPAPNITCKDSSGNETH 161

Db 20 NVDCNELQMMQFMCPDPARGQIDPKTQQLAGCTREGARVWCIAANEINCFE-TGNAI- 77

QY 162 FTGNEVGFPPISCRWNVNGSYKVAVALSLFLGLWGLADREYLYGYPALGLLKFCTVFCGI 221

Db 78 -----FTREVCKWTNGVHLDITLLSVLFMGFGVDREYLYGYPGIGLLKFCITLGMEL 130

QY 222 GSLIDFILSMQIVGSDGSSYIIDYGTSLTSLI-NETFR 263

Db 131 GOLIDIVLIALQVVGPDGSAIVIPYTGAGIHIVRSNDNTTYR 172

RESULT 8

AAU97631

ID AAU97631 standard; Protein: 100 AA.

XX AC AAU97631;

XX DT 13-AUG-2002 (first entry)

XX DE RNA polymerase II subunit 11 protein.

XX KW RNA polymerase II subunit 11; cancer; HIV; infection; human immunodeficiency virus.

XX OS Unidentified.

XX PN CN1331300-A.

XX PD 16-JAN-2002.

XX PF 30-JUN-2000; 2000CN-0116963.

XX PR 30-JUN-2000; 2000CN-0116963.

XX PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.

XX PI Mao Y, Xie Y;

XX DR WPI: 2002-340664/38.

XX DR N-PSDB: ABK52558.

XX PT Polypeptide-RNA polymerase II subunit 11 and polynucleotide for coding it.

XX PS Claim 1; Page 29; 32pp; Chinese.

CC CC This invention relates to the DNA and protein sequences of a novel polypeptide-RNA polymerase II subunit 11 protein. The invention also comprises a process for preparing the polypeptide of the invention by DNA recombination, the application of the polypeptide in treating diseases such as cancer, human immunodeficiency virus (HIV) infection, etc, the antagonist of the polypeptide and its medical action, and the application of the said polynucleotide are disclosed. The present sequence represents the RNA polymerase II subunit 11 protein of the invention.

XX SQ Sequence 100 AA;

Query Match 20.4%; Score 293; DB 23; Length 100;

Best Local Similarity 58.2%; Pred. No. 1.4e-22;

Matches 55; Conservative 1; Mismatches 0; Indels 0; Gaps 6;

QY 63 MAANWPSGSPAPENVTARLVGLVFWVSVTGPGAVATSGAGFESLKCEDLVKGQY 118

Db 1 MAANWPSGSPAPDAVTARLVGLVFWVSVTGPGAVATSGAGFESLKCEDLVKGQY 56

RESULT 9

ABR47818

ID ABR47818 standard; Protein: 221 AA.

XX AC ABR47818;

XX DT 12-JUN-2003 (first entry)

XX DE Human secreted protein, SEQ ID 709.

XX KW Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cytostatic; vulnerability; antiinflammatory; nootropic; neuroprotective; antiparkinsonian; gene therapy; human; cardiovascular disorder.

XX OS Homo sapiens.

XX PN WO200295010-A2.

PD 28-NOV-2002.
 XX
 PF 19-MAR-2002; 2002WO-US09785.
 XX
 PR 21-MAR-2001; 2001US-277340P.
 PR 19-JUL-2001; 2001US-306171P.
 PR 13-NOV-2001; 2001US-331287P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2003-129429/12.
 XX
 PT Novel human secreted proteins, useful for detecting, preventing,
 PT diagnosing, prognosticating, treating and/or ameliorating
 PT cardiovascular disorders such as arrhythmia -
 XX
 PS Claim 13; SEQ ID 709; 1881pp; English.
 XX
 CC The present invention relates to novel human secreted proteins
 CC (AB47633-ABR48145) and their coding sequences (ACC50344-ACC50856). The
 CC proteins and their coding sequences are useful for the preparation of a
 CC diagnostic or pharmaceutical composition for diagnosing or treating a
 CC cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,
 CC coronary arteriosclerosis and myocardial ischemia), neural disorders,
 CC immune system disorders, muscular disorders, reproductive disorders,
 CC gastrointestinal disorders, pulmonary disorders, renal disorders,
 CC proliferative disorders and/or cancerous diseases and conditions, for
 CC wound healing and epithelial cell proliferation, to treat inflammation or
 CC infection, for treating thrombosis and arteriosclerosis, for treating or
 CC preventing neural damage which occurs in neuronal disorders or
 CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's
 CC disease, to enhance bone and periodontal regeneration and aid in tissue
 CC transplants or bone grafts, to prevent skin aging or hair loss, to
 CC stimulate growth and differentiation of haematopoietic cells and bone
 CC marrow cells when used in combination with other cytokines, to maintain
 CC organs before transplantation or for supporting cell culture of primary
 CC tissues, to increase or decrease differentiation or proliferation of
 CC embryonic stem cells, or to modulate mammalian characteristics or
 CC metabolism.
 CC Note: the sequence data for this patent was published in electronic
 CC format and is available from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 221 AA;
 Query Match 14.0%; Score 201; EB 24; Length 221;
 Best Local Similarity 45.7%; Pred. No. 1.9e-12;
 Matches 53; Conservative 12; Mismatches 39; Indels 12; Gaps 5;
 QY 135 CTNYTA--HVSC----FPAPNITCKDSSGNE:HFTGNEVGFFKPISCRNYNGYSKYKAVA 183
 Db 112 CTNISTCMTVSCPRORYPA-NCTVRD----HVHCLGNRT-FPKMLYCNKTCGYKASTALA 165
 QY 189 LSLFLGWLGAADRFYLGYPALGLLCKTFCGCGISGLIDFLISMQIVGSPGSSSY 244
 Db 166 LSLITLGGFGADRFYLGQWKEGLKFLSFGGLGIWTLIDVLLIGVGVGPADGSLYI 221
 RESULT 10
 ABR00112
 ID ABR00112 standard; Protein: 221 AA.
 XX
 AC ABR00112;
 XX
 DT 03-APR-2003 (first entry)
 XX
 DE Human gene 102 encoded secreted protein HMEED17, SEQ ID NO:401.
 XX
 KW Human; secreted protein; digestive disorder; gastrointestinal disorder;
 KW mouth; oesophagus; stomach; small intestine; large intestine; liver;
 KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;

KW immune disorder; inflammation; infection; wound healing; drug screening;
 KW chromosome identification; chromosome mapping; cytostatic;
 KW antiinflammatory; immunosuppressive; vulnery; gene therapy.
 XX
 OS Homo sapiens.
 PN WO200276488-A1.
 XX
 PD 03-OCT-2002.
 XX
 PF 19-MAR-2002; 2002WO-US08276.
 XX
 PR 21-MAR-2001; 2001US-277340P.
 PR 19-JUL-2001; 2001US-306171P.
 PR 13-NOV-2001; 2001US-331287P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2003-029900/02.
 DR N-PSDB; ABZ71291.
 XX
 PT New human secreted proteins and nucleic acids, useful for detecting,
 PT preventing, diagnosing, prognosticating, treating and/or ameliorating
 PT e.g. gastrointestinal diseases and disorders, or cancers -
 XX
 PS Claim 13; Page 1007; 1216pp; English.
 XX
 CC ABZ71190-ABZ71478 represent cDNAs corresponding to 178 human secreted
 CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.
 CC ABZ71479-ABZ71540 represent human secreted protein genomic fragments. The
 CC invention also encompasses antibodies specific for the secreted proteins
 CC the use of the secreted proteins in drug screening, and recombinant
 CC vectors and host cells comprising a nucleic acid of the invention. The
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody
 CC fragments specific for the secreted proteins, and modulators of protein
 CC activity are useful for diagnosing, treating, ameliorating or preventing
 CC digestive disorders. Such conditions include disorders of the mouth,
 CC oesophagus, stomach, small intestine, large intestine, liver, biliary
 CC tract and pancreas, and include cancers of these organs and tissues. The
 CC secreted proteins and their nucleic acids may also be used in the
 CC treatment of immune disorders, inflammation, infection,
 CC hyperproliferative disorders, and to promote wound healing. Nucleic acids
 CC of the invention may be used for chromosome identification, chromosome
 CC mapping, in gene therapy, for identifying individuals from minute
 CC biological samples, as hybridisation probes, and as molecular weight
 CC markers. The present sequence represents a human secreted protein of the
 CC invention.
 XX
 SQ Sequence 221 AA;
 Query Match 14.0%; Score 201; DB 24; Length 221;
 Best Local Similarity 45.7%; Pred. No. 1.9e-12;
 Matches 53; Conservative 12; Mismatches 39; Indels 12; Gaps 5;
 QY 135 CTNYTA--HVSC----FPAPNITCKDSSGNE:HFTGNEVGFFKPISCRNYNGYSKYKAVA 183
 Db 112 CTNISTCMTVSCPRORYPA-NCTVRD----HVHCLGNRT-FPKMLYCNKTCGYKASTALA 165
 QY 189 LSLFLGWLGAADRFYLGYPALGLLCKTFCGCGISGLIDFLISMQIVGSPGSSSY 244
 Db 166 LSLITLGGFGADRFYLGQWKEGLKFLSFGGLGIWTLIDVLLIGVGVGPADGSLYI 221
 RESULT 11
 AAY27658
 ID AAY27658 standard; Protein: 222 AA.
 XX
 AC AAY27658;
 XX
 DT 30-JUL-1999 (first entry)
 XX

Db 112 CTNSTSCMTVSCPRORYPA-NCTVRD----HVHCLGNRT-PPKMLYCWNVTGGYKWSIALA 165
 QY 189 LSLFLGNLGAADRFYLCYPAJGLLKCTVFCGIGSLIDFILSMQIVGPSGSSVI 244
 XX 112 CTNSTSCMTVSCPRORYPA-NCTVRD----HVHCLGNRT-PPKMLYCWNVTGGYKWSIALA 165
 Db 166 LSLTLGGFGADRFYLCQWREGGLKLFSGGLGIWTLIDVLLIGVGVGPADGSLVI 221

RESULT 13

ABR48493
 ID ABR48493 standard: Protein: 221 AA.

AC ABR48493:

XX 13-JUN-2003 (first entry)

XX Human Anyloid Apoptotic Receptor (AAR); protein.

XX Human: GENSET; therapeutic; therapy.

XX Homo sapiens.

XX WO200204864-A2.

XX 28-NOV-2002.

XX 06-AUG-2001; 2001WO-IB01715.

XX 25-MAY-2001; 2001US-293574P.

XX 15-JUN-2001; 2001US-298698P.

XX 29-JUN-2001; 2001US-302277P.

XX 13-JUL-2001; 2001US-305456P.

XX (GEST) GENSET.

XX Bejanin S, Tanaka H;

XX WPI; 2003-129412/12.

XX N-PSDB; ACC51100.

XX New GENSET polynucleotides and polypeptides, useful for preparing a

XX composition for treating GENSET-related disorders and as reagents in

XX assays to quantitatively determined levels of GENSET expression in

XX biological samples .

XX Claim 2; Page 475-476; 505pp; English.

XX The present invention relates to novel human GENSET coding sequences

XX (ACC51050-ACC51115) and proteins (ABR48453-ABR48508). The GENSET

XX sequences are useful for preparing a composition for treating

XX GENSET-related disorders. They can also be used as markers for tissues in

XX which the corresponding protein is preferentially expressed, as molecular

XX weight markers on Southern gels, as chromosome markers or tags to

XX identify chromosomes, and as reagents in assays to quantitatively

XX determined levels of GENSET expression in biological samples.

XX Sequence 221 AA;

XX Query Match 13.9%; Score 200; DB 24; Length 221;

XX Best Local Similarity 45.7%; Pred. No. 2.4e-12;

XX Matches 53; Conservative 12; Mismatches 39; Indels 12; Gaps 5;

AB08641 standard; Protein: 247 AA.

AC ABR08641;

DI 02-JAN-2001 (first entry)

DE A human TANGO 223 polypeptide.

XX TANGO 201; TANGO 223; secreted protein; transmembrane protein;
 KW pro-liferative disorder; neoplasm; tumour; pancreatic disorder; cirrhosis
 KW pancreatitis; thyroiditis; goiter; Graves' disease; gastric disorder;
 KW gastritis; tumour; placental disorder; placentitis; spontaneous abortion
 KW pulmonary disorder; atelectasis; edema; Goodpasture's syndrome;
 KW muscular dystrophy; cardiovascular disorder; ischemic heart disease;
 KW congenital heart disease; cerebral edema; cerebrovascular disease;
 KW hepatic disorder; jaundice; hepatitis; leukocytic disorder; hyperplasia;
 KW leukopenia; leukocytosis; malignant lymphoma; prostate disorder;
 KW inflammatory disease.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..29

FT Modified-site 28..31 /note= "signal peptide"

FT Modified-site 29..34 /note= "casein kinase II phosphorylation site"

FT Protein 30..247 /note= "myristoylation site"

FT Domain 30..215 /note= "mature protein"

FT Modified-site 39..42 /note= "extracellular domain"

FT Modified-site 39..41 /note= "casein kinase II phosphorylation site"

FT Modified-site 39..90 /note= "protein kinase C phosphorylation site"

FT Modified-site 107..110 /note= "N-glycosylation site"

FT Modified-site 115..117 /note= "casein kinase II phosphorylation site"

FT Modified-site 122..125 /note= "protein kinase C phosphorylation site"

FT Modified-site 124..126 /note= "N-glycosylation site"

FT Modified-site 134..137 /note= "protein kinase C phosphorylation site"

FT Modified-site 140..143 /note= "casein kinase II phosphorylation site"

FT Modified-site 157..160 /note= "N-glycosylation site"

FT Modified-site 159..161 /note= "N-glycosylation site"

FT Modified-site 159..162 /note= "protein kinase C phosphorylation site"

FT Modified-site 169..172 /note= "casein kinase II phosphorylation site"

FT Modified-site 179..182 /note= "N-glycosylation site"

FT Modified-site 197..202 /note= "N-glycosylation site"

FT Domain 216..238 /note= "myristoylation site"

FT Modified-site 226..229 /note= "transmembrane domain"

FT Domain 239..247 /note= "casein kinase II phosphorylation site"

XX WO200050442-A2.

XX 31-AUG-2000.

XX

RESULT 14

AB08641

```

PF 25-FEB-2000; 2000WO-US04784.
XX
XX
PR 26-FEB-1999; 99US-0259388.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX McCarthy SA;
XX
XX WPI: 2000-572066/53.
DR N-PSDB; AAA64408; AAA64409.
XX
XX Novel nucleic acid molecule encoding secreted or transmembrane protein
PT useful for identifying modulators and for diagnosing and treating
PT pancreatic, cardiovascular, liver and pituitary disorders -
XX
XX Claim 9; Fig 8; 176pp; English.
XX
XX The present sequence represents TANGO 223 polypeptide. The specification
CC also describes a TANGO 201 polypeptide. These polypeptides are secreted
CC or transmembrane proteins. Human TANGO 201 and 223 nucleic acids,
CC proteins and their modulators are useful for treating proliferative
CC disorders e.g. neoplasms or tumours, pancreatic disorders
CC (e.g. pancreatitis), disorders of the adrenal cortex, adrenal medulla,
CC thyroid gland (e.g. thyroiditis), goiter, Graves' disease, gastric
CC disorders (e.g. gastritis or tumours), placental disorders (e.g.
CC placentalis or spontaneous abortion), pulmonary disorders (e.g.
CC atelectasis), edema, Goodpasture's syndrome, disorders of the skeletal
CC muscle (e.g. muscular dystrophy), cardiovascular disorders (e.g.
CC ischemic heart disease and congenital heart disease), disorders of the
CC brain (e.g. cerebral edema), cerebrovascular disease and to treat
CC injury or trauma to the brain. They are also useful for treating
CC hepatic disorders (e.g. jaundice, hepatitis, cirrhosis or malignant
CC tumours), renal, testicular, intestinal disorders. TANGO 223
CC polynucleotides are also useful for treating leukocytic disorders (e.g.
CC leukopenias, leukocytosis and malignant lymphomas) and prostate
CC disorders (e.g. inflammatory diseases, hyperplasia or tumours).
XX
XX Sequence 247 AA;
XX
XX Query Match 13.9%; Score 200; DB 21; Length 247;
XX Best Local Similarity 45.7%; Pred. No. 2.8e-12;
XX Matches 53; Conservative 12; Mismatches 39; Indels 12; Gaps 5;
XX
QY 135 CTNYTA--HVSC---FPAPNITCKDSSGNETHFTGNEVGFKPISCRNVNGYSYKVAVA 188
11 13 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 138 CTNSTSMTVSCPRORYPA-NCTVRD---HVHCLGNRT-FPKMLYCNWTGGYKXSTALA 191
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
QY 189 LSLFLGWLGAADRYLGYPALGLLKFTCTVCGCGISLIDFILISMQIVGSPDGSYYI 244
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 192 LSITLGGGADRYLQWREGGLKLFSGGLGIWTLIDVLLIGVGVGPDGSLYI 247
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11

RESULT 15
AAB08656
ID AAB08656 standard; Protein: 247 AA.
XX
XX AAB08656;
XX
XX 02-JAN-2001 (first entry)
XX
XX A human TANGO 223 variant polypeptide.
XX
XX TANGO 201; secreted protein; transmembrane protein;
XX proliferative disorder; neoplasms; tumour; pancreatic disorder; cirrhosis;
XX pancreatitis; thyroiditis; goiter; Graves' disease; gastric disorder;
XX gastritis; tumour; placental disorder; placentalis; spontaneous abortion;
XX pulmonary disorder; atelectasis; edema; Goodpasture's syndrome;
XX muscular dystrophy; cardiovascular disorder; ischemic heart disease;
XX congenital heart disease; cerebral edema; cerebrovascular disease;
XX hepatic disorder; jaundice; hepatitis; leukocytic disorder; hyperplasia;
XX leukopenia; leukocytosis; malignant lymphoma; prostate disorder;
XX inflammatory disease.
XX

```

```

OS Homo sapiens.
XX
XX WO2000050442-A2.
XX
XX 31-AUG-2000.
XX
XX 25-FEB-2000; 2000WO-US04784.
XX
XX 26-FEB-1999; 99US-0259388.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX McCarthy SA;
XX
XX WPI: 2000-572066/53.
DR N-PSDB; AAA64424.
XX
XX Novel nucleic acid molecule encoding secreted or transmembrane protein
PT useful for identifying modulators and for diagnosing and treating
PT pancreatic, cardiovascular, liver and pituitary disorders -
XX
XX Claim 9; Page 169-170; 176pp; English.
XX
XX The present sequence represents a TANGO 223 variant. The specification
CC also describes a TANGO 201 polypeptide. These polypeptides are secreted
CC or transmembrane proteins. Human TANGO 201 and 223 nucleic acids,
CC proteins and their modulators are useful for treating proliferative
CC disorders e.g. neoplasms or tumours, pancreatic disorders
CC (e.g. pancreatitis), disorders of the adrenal cortex, adrenal medulla,
CC thyroid gland (e.g. thyroiditis), goiter, Graves' disease, gastric
CC disorders (e.g. gastritis or tumours), placental disorders (e.g.
CC placentalis or spontaneous abortion), pulmonary disorders (e.g.
CC atelectasis), edema, Goodpasture's syndrome, disorders of the skeletal
CC muscle (e.g. muscular dystrophy), cardiovascular disorders (e.g.
CC ischemic heart disease and congenital heart disease), disorders of the
CC brain (e.g. cerebral edema), cerebrovascular disease and to treat
CC injury or trauma to the brain. They are also useful for treating
CC hepatic disorders (e.g. jaundice, hepatitis, cirrhosis or malignant
CC tumours), renal, testicular, intestinal disorders. TANGO 223
CC polynucleotides are also useful for treating leukocytic disorders (e.g.
CC leukopenias, leukocytosis and malignant lymphomas) and prostate
CC disorders (e.g. inflammatory diseases, hyperplasia or tumours).
XX
XX Sequence 247 AA;
XX
XX Query Match 13.9%; Score 200; DB 21; Length 247;
XX Best Local Similarity 45.7%; Pred. No. 2.8e-12;
XX Matches 53; Conservative 12; Mismatches 39; Indels 12; Gaps 5;
XX
QY 135 CTNYTA--HVSC---FPAPNITCKDSSGNETHFTGNEVGFKPISCRNVNGYSYKVAVA 188
11 13 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 138 CTNSTSMTVSCPRORYPA-NCTVRD---HVHCLGNRT-FPKMLYCNWTGGYKXSTALA 191
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
QY 189 LSLFLGWLGAADRYLGYPALGLLKFTCTVCGCGISLIDFILISMQIVGSPDGSYYI 244
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 192 LSITLGGGADRYLQWREGGLKLFSGGLGIWTLIDVLLIGVGVGPDGSLYI 247
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11

```

Search completed: September 26, 2003, 17:25:14
Job time : 89 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw mode:

Run on: September 26, 2003, 17:25:22 : Search time 30 seconds
(without alignments)
379,387 Million cell updates/sec

Title: US-09-852-100A-2

Perfect score: 1439

Sequence: 1 MHILKGSNPVTPRAHGQKNI.....TLRLSLINETFRKIQIYP 269

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued_Patents_AA:*
- 1: /cgn2.6/ptodata/1/1aa/5A-COMB.pep:*
 - 2: /cgn2.6/ptodata/1/1aa/5B-COMB.pep:*
 - 3: /cgn2.6/ptodata/1/1aa/6A-COMB.pep:*
 - 4: /cgn2.6/ptodata/1/1aa/6B-COMB.pep:*
 - 5: /cgn2.6/ptodata/1/1aa/PCTUS-COMB.pep:*
 - 6: /cgn2.6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed. and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %		Match	Length	DB ID	Description:
		Match	Length				
1	85	5.9	310	2	US-08-414-657D-45	Sequence 45, Appl	Sequence 45, Appl
2	85	5.9	338	2	US-08-414-657D-42	Sequence 42, Appl	Sequence 42, Appl
3	85	5.9	338	2	US-08-414-657D-43	Sequence 43, Appl	Sequence 43, Appl
4	85	5.9	338	4	US-09-135-080-4	Sequence 4, Appl	Sequence 4, Appl
5	84	5.8	258	4	US-09-328-352-4253	Sequence 4253, Ap	Sequence 4253, Ap
6	83	5.8	764	2	US-08-177-109A-2	Sequence 2, Appl	Sequence 2, Appl
7	83	5.8	764	2	US-08-687-706-2	Sequence 2, Appl	Sequence 2, Appl
8	81	5.6	797	3	US-09-182-728A-2	Sequence 2, Appl	Sequence 2, Appl
9	81	5.6	797	4	US-09-795-232-2	Sequence 2, Appl	Sequence 2, Appl
10	80.5	5.6	150	4	US-09-252-991A-16958	Sequence 16958, A	Sequence 16958, A
11	80.5	5.6	304	2	US-08-414-657D-44	Sequence 44, Appl	Sequence 44, Appl
12	80.5	5.6	325	2	US-08-414-657D-2	Sequence 2, Appl	Sequence 2, Appl
13	80.5	5.6	325	2	US-08-414-657D-41	Sequence 41, Appl	Sequence 41, Appl
14	80.5	5.6	325	4	US-09-135-080-2	Sequence 2, Appl	Sequence 2, Appl
15	80.5	5.6	456	3	US-09-058-389A-4	Sequence 4, Appl	Sequence 4, Appl
16	80.5	5.6	456	4	US-09-611-781-4	Sequence 4, Appl	Sequence 4, Appl
17	80.5	5.6	1237	1	US-08-241-853-2	Sequence 2, Appl	Sequence 2, Appl
18	80.5	5.6	1237	4	US-08-850-917-2	Sequence 2, Appl	Sequence 2, Appl
19	80	5.6	328	4	US-09-253-316-28	Sequence 28, Appl	Sequence 28, Appl
20	80	5.6	660	3	US-09-181-706-8	Sequence 8, Appl	Sequence 8, Appl
21	80	5.6	660	3	US-09-458-791-8	Sequence 8, Appl	Sequence 8, Appl
22	80	5.6	660	3	US-09-459-066-8	Sequence 6, Appl	Sequence 6, Appl
23	80	5.6	660	4	US-09-459-066-8	Sequence 8, Appl	Sequence 8, Appl
24	79.5	5.5	354	2	US-08-700-013B-9	Sequence 9, Appl	Sequence 9, Appl
25	79.5	5.5	797	2	US-08-700-013B-21	Sequence 21, Appl	Sequence 21, Appl
26	79.5	5.5	927	4	US-09-328-352-7922	Sequence 7922, Ap	Sequence 7922, Ap
27	79	5.5	797	2	US-08-700-013B-19	Sequence 19, Appl	Sequence 19, Appl

Sequence 124, Appl
Sequence 19777, A
Sequence 60, Appl
Sequence 8, Appl
Sequence 2, Appl
Sequence 31718, A
Sequence 26357, A
Sequence 120, Appl
Sequence 122, Appl
Sequence 4, Appl
Sequence 26323, A
Sequence 2, Appl
Sequence 2, Appl
Sequence 21426, A
Sequence 3, Appl
Sequence 37, Appl
Sequence 11, Appl
Sequence 42, Appl

ALIGNMENTS

RESULT 1
US-08-414-657D-45
: Sequence 45, Application US/08414657D
: Patent No. 5861283
: GENERAL INFORMATION:
: APPLICANT: Iovitt, Pat
: APPLICANT: Pimenta, Aurea
: APPLICANT: Fischer, Itzhak
: APPLICANT: Zhukareva, Victoria
: TITLE OF INVENTION: Limbic System-Associated Membrane
: TITLE OF INVENTION: Protein and DNA
: NUMBER OF SEQUENCES: 60
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dechert Price & Rhoads
: STREET: 997 Lenox Drive, Building 3, Suite 210
: CITY: Lawrenceville
: STATE: NJ
: COUNTRY: USA
: ZIP: 08543
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA: US/08/414.657D
: FILING DATE: 31-MAR-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Bloom, Allen
: REGISTRATION NUMBER: 29,135
: REFERENCE/DOCKET NUMBER: 317743-102
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 609-520-3214
: TELEFAX: 609-520-3259
: TELEX:
: INFORMATION FOR SEQ ID NO: 45:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 310 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
US-08-414-657D-45

Query Match 5.9%; Score 85; DB 2; Length 310;
Best Local Similarity 27.7%; Pred. NO. 1.7;
Matches 36; Conservative 15; Mismatches 47; Indels 32; Gaps 7;

Db 2 GSNLSP-----QCLMPPFILLLSGGVTTPWSLAQPQCSLEGVIEKGGSPRLQEG 55
Qy 61 SKMAAMPSPG--PSAPEAVTARLVGLKFWVITGPMCAVATS-----AGGEESLKC-- 110
Db 56 QALEYVCPGPGYPVQVTRC-----STGWSSTLKTDQKTVKRAECRAHCPR 105
Qy 111 -EDLKVGQYICKPKINDATQEPVNC-TNYTAHVSCFPAPNITCKDSS--GNETHFTGNE 166
Db 106 PHDFENGEMPRSPYVNVSDISFHCYDGYTLRGS-----NRTCOVNGRWSGQTALCDNG 161
Qy 167 VGPFK-----PISCRNVNGYSYKV 185
Db 162 AGYCSNPGIPGTRKV-GSQYRL 183

RESULT 7

US-08-687-706-2
: Sequence 2, Application US/04687706
: Patent No. 5928892
: GENERAL INFORMATION:
: APPLICANT: Dennis E. Hourcade and Teresa J. Oakesby
: TITLE OF INVENTION: MODIFIED COMPLEMENT PROTEASES
: NUMBER OF SEQUENCES: 62
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Patrea L. Pabst
: STREET: 2800 One Atlantic Center
: CITY: Atlanta
: STATE: Georgia
: COUNTRY: USA
: ZIP: 30309-3450
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/687,706
: FILING DATE: 26-JUL-1996
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/177,109
: FILING DATE: 03-JAN-1994
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Pabst, Patrea L.
: REGISTRATION NUMBER: 31,284
: REFERENCE/DOCKET NUMBER: WU 107 DIV
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (404) 873-8794
: TELEFAX: (404) 873-8795
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 764 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHEICAL: NO
US-08-687-706-2

Query Match 5.8%; Score 83; DB 2; Length 764;
Best Local Similarity 24.1%; Pred. No. 10;
Matches 49; Conservative 21; Mismatches 7; Indels 62; Gaps 12;
Qy 24 GTGLPMRGPKNLALPFLPLGCG-----GSGSGFKSV----- 60
Db 2 GSNLSP-----QCLMPPFILLLSGGVTTPWSLAQPQCSLEGVIEKGGSPRLQEG 55
Qy 61 SKMAAMPSPG--PSAPEAVTARLVGLKFWVITGPMCAVATS-----AGGEESLKC-- 110
Db 56 QALEYVCPGPGYPVQVTRC-----STGWSSTLKTDQKTVKRAECRAHCPR 105
Qy 111 -EDLKVGQYICKPKINDATQEPVNC-TNYTAHVSCFPAPNITCKDSS--GNETHFTGNE 166

Db 106 PHDFENGEMPRSPYVNVSDISFHCYDGYTLRGS-----NRTCOVNGRWSGQTALCDNG 161
Qy 167 VGPFK-----PISCRNVNGYSYKV 185
Db 162 AGYCSNPGIPGTRKV-GSQYRL 183

RESULT 8

US-09-182-728A-2
: Sequence 2, Application US/09182728A
: Patent No. 6238883
: GENERAL INFORMATION:
: APPLICANT: BROWN, ANTHONY
: APPLICANT: CHAPMAN, CONRAD GERALD
: APPLICANT: GLOGER, ISRAEL SIMON
: APPLICANT: EVANS, JOANNE RACHEL
: APPLICANT: CAIRNS, WILLIAM
: APPLICANT: HERDON, HUGH
: TITLE OF INVENTION: NOVEL COMPOUNDS
: FILE REFERENCE: GP-30176
: CURRENT APPLICATION NUMBER: US/09/182,728A
: CURRENT FILING DATE: 1998-10-29
: PRIOR APPLICATION NUMBER: 9818890.7
: PRIOR FILING DATE: 1998-08-28
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 2
: LENGTH: 797
: TYPE: PRT
: ORGANISM: HOMO SAPIENS
US-09-182-728A-2
Query Match 5.6%; Score 81; DB 3; Length 797;
Best Local Similarity 23.9%; Pred. No. 17;
Matches 47; Conservative 32; Mismatches 80; Indels 38; Gaps 21;
Qy 87 FSVTTGPMGAVATSAGGEESLKCED---LKVQGYICKD-PKINDATQEPVNCNTYAHV 342
Db 302 FVSVL--PMGSCNNPMTPE---CKDKTKLLDSCVSDHPKI-----QIKNSTFCM 348
Qy 143 SCPPAPNITCKDSSGNETHFTGNEVGFPRPISCRNVNGYSY-----KVAVALSFJGWLGA 198
Db 349 TAYPNVTMWNFTSQANKTFVSGSE-EYFKYVLKISAGIEYGEIRMPALCLFLAW-- 405
Qy 199 DREYLGYPAL-----GLLKFCCTGFCGIGSLIDFILISMQIVGSPSGSYLIDYVGR 251
Db 406 ----IVYASLAKGIK*SGKVYFTATFPYV-VLVILLIRGVTLPGAGAGIWFITPKWEK 460
Qy 252 LIRLSITNEIFRKTQY 268
Db 461 LTNATVWKDA--ATOIF 475
RESULT 9
US-09-795-232-2
: Sequence 2, Application US/09795232
: Patent No. 6426405
: GENERAL INFORMATION:
: APPLICANT: Anthony M. Brown
: APPLICANT: Conrad Gerald Chapman
: APPLICANT: Israel Simon Gloger
: APPLICANT: Joanne Rachel Evans
: APPLICANT: William Cairns
: APPLICANT: Hugh Jonathan Herdon
: TITLE OF INVENTION: NOVEL COMPOUNDS
: FILE REFERENCE: GP-30176-D1
: CURRENT APPLICATION NUMBER: US/09/795,232
: CURRENT FILING DATE: 2001-02-28
: PRIOR APPLICATION NUMBER: 09/182,728
: PRIOR FILING DATE: 1998-10-29
: PRIOR APPLICATION NUMBER: 9818890.7
: PRIOR FILING DATE: 1998-08-28

```

; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 797
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-795-232-2

Query Match.          5.6%; Score 81; DB 4; Length 797;
Best Local Similarity 23.9%; Pred. No. 17;
Matches 47; Conservative 32; Mismatches 80; Indels 38; Gaps 11;

QY 87 FVSVTTGPGCAVATISAGGESLACED--LKVGQYICKD-PKINDATQEPVNCNTYTAHV 142
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 302 FVSVL--PMGSCNPNTPB---CKDKLLKSSCVISHPKI-----QTKNSTFCM 348

QY 143 SCFPAPNITCKSSGNETHTFTGNEVGFFPKPISCRNVNGYSY---KVAVALSFLGLG 196
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 349 TAYPNVTWVNFTSQANKTFVSGSE-EYFKYFVLKISAGIEYPGEIRWPLALCIFLAKV-- 405

QY 199 DRYLGYPAL-----GLLKFTVGFPGGIGSLIDFILISKQIVGSDGSSYLDYFGR 251
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 406 ----IVYASLAKGIKTSKVVYFTATPPYV-VIVILLIRGVTLPGAGAGIWFITPKWEK 460

QY 252 LIRLSIINETFRKTQLY 268
   || || || || || || || || || || || || || || || || || || || ||
Db 461 LIRATVWKDA--ATQIF 475

RESULT 10
US-09-252-991A-16958
; Sequence 16958, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16958
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16958

Query Match          5.6%; Score 80.5; DB 4; Length 150;
Best Local Similarity 40.0%; Pred. No. 1.8;
Matches 20; Conservative 7; Mismatches 22; Indels 1; Gaps 1;

QY 181 YSKVAVALSFLGLGWLGAADRFYLCYPALGLKFCUTV;FCGIGSLIDFIL; 230
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 24 HSAKIGLLMIF-GFTGSHRFRY;GKPIITGIIWFTFGIFGLFIIGIIDLFI 72

RESULT 11
US-08-414-657D-44
; Sequence 44, Application US/08414657D
; Patent No. 5861283
; GENERAL INFORMATION:
; APPLICANT: Levitt, Pat
; APPLICANT: Pimenta, Aurea
; APPLICANT: Fischer, Itzhak
; APPLICANT: Zhukareva, Victoria
; TITLE OF INVENTION: Limbic System-Associated Membrane
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 997 Lenox Drive, Building 3, Suite 210
; CITY: Lawrenceville
; STATE: NJ
; COUNTRY: USA
; ZIP: 08543
; COMPUTER READABLE FORM:
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/414,657D
; FILING DATE: 31-MAR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION: 435
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bloom, Allen
; REGISTRATION NUMBER: 29,135
; REFERENCE/DOCKET NUMBER: 317743-102
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 609-520-3214
; TELEFAX: 609-520-3259
; TELEX:
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 304 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-414-657D-44

Query Match          5.6%; Score 80.5; DB 2; Length 304;
Best Local Similarity 29.9%; Pred. No. 4.8;
Matches 32; Conservative 14; Mismatches 40; Indels 21; Gaps 6;

QY 201 SAGGESLACEDLKVG-----QYICKDPKINDATQEPVNCNTYTAHVSCFPAPNITCKDSS 116
   : : : ||||| : : : : : : : : : : : : : : : : : : : : : : :
Db 202 TTGRQASLKCEASAVPADPFEWYRDDTRINSANGLEIKSTE---GQSSLTVTNVT-EEHY 217
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 157 GNETHTFTGNEVG-----FFKPISCRNVNGYSYKVAVALSFLGLWL 196
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 258 GNYTCVAAKLGVTNANSLVLFHFGSVRGNG-SISIAVPL-----WL 298
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
US-08-414-657D-2
; Sequence 2, Application US/08414657D
; Patent No. 5861283
; GENERAL INFORMATION:
; APPLICANT: Levitt, Pat
; APPLICANT: Pimenta, Aurea
; APPLICANT: Fischer, Itzhak
; APPLICANT: Zhukareva, Victoria
; TITLE OF INVENTION: Limbic System-Associated Membrane
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 997 Lenox Drive, Building 3, Suite 210
; CITY: Lawrenceville
; STATE: NJ
; COUNTRY: USA
; ZIP: 08543
; COMPUTER READABLE FORM:
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/414,657D

```

;; FILING DATE: 31-MAR-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Bloom, Allen
;; REGISTRATION NUMBER: 29,135
;; REFERENCE/DOCKET NUMBER: 317743-102
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 609-520-3214
;; TELEFAX: 609-520-3259
;; TELEX:

;; INFORMATION FOR SEQ ID NO: 2:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 325 amino acids

;; TYPE: amino acid

;; STRANDEDNESS: single

;; TOPOLOGY: linear

;; MOLECULE TYPE: protein

;; FRAGMENT TYPE: internal

US-08-414-657D-2

Query Match 5.6%; Score 80.5; DB 2; Length 325;
Best Local Similarity 29.9%; Pred. No. 5.3;
Matches 32; Conservative 14; Mismatches 40; Indels 21; Gaps 6;

Qy 101 SAGGEESLKCEDLVG-----QYICKDPKINDATQEPVNTNVTARHVSCEFAPNITCKDSS 156
Db 223 TTGRQASLKCEASAVPAPDFEYRDDTRINSANGLEIKSTE---GOSSLVTNVT-EEHY 273
Qy 157 GNEHTFTONEVG-----FFKPISCRNVNGYSYKVAVALSLFLGWL 196
Db 279 GNYTCVAANKLGVTNASLVLPFGSVRGING-SISLAVPL-----WL 319

RESULT 13

US-08-414-657D-41
; Sequence 41, Application US/08414657D
; Patent No. 5861283

;; GENERAL INFORMATION:

;; APPLICANT: Lovitt, Pat

;; APPLICANT: Pimenta, Aurea

;; APPLICANT: Fischer, Itzhak

;; APPLICANT: Zhukareva, Victoria

;; TITLE OF INVENTION: Limbic System-Associated Membrane

;; TITLE OF INVENTION: Protein and DNA

;; NUMBER OF SEQUENCES: 60

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: Dechert Price & Rhoads

;; STREET: 997 Lenox Drive, Building 3, Suite 210

;; CITY: Lawrenceville

;; STATE: NJ

;; COUNTRY: USA

;; ZIP: 08543

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Diskette

;; COMPUTER: IBM Compatible

;; OPERATING SYSTEM: DOS

;; SOFTWARE: FastSEQ for Windows Version 2.0

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/414,657D

;; FILING DATE: 31-MAR-1995

;; CLASSIFICATION: 435

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER:

;; FILING DATE:

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Bloom, Allen

;; REGISTRATION NUMBER: 29,135

;; REFERENCE/DOCKET NUMBER: 317743-102

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: 609-520-3214

;; TELEFAX: 609-520-3259

;; TELEX:

;; INFORMATION FOR SEQ ID NO: 41:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 325 amino acids

;; TYPE: amino acid

;; STRANDEDNESS: single

;; TOPOLOGY: linear

US-08-414-657D-41

Query Match 5.6%; Score 80.5; DB 2; Length 325;
Best Local Similarity 29.9%; Pred. No. 5.3;
Matches 32; Conservative 14; Mismatches 40; Indels 21; Gaps 6;

Qy 101 SAGGEESLKCEDLVG-----QYICKDPKINDATQEPVNTNVTARHVSCEFAPNITCKDSS 156
Db 223 TTGRQASLKCEASAVPAPDFEYRDDTRINSANGLEIKSTE---GOSSLVTNVT-EEHY 273
Qy 157 GNEHTFTONEVG-----FFKPISCRNVNGYSYKVAVALSLFLGWL 196
Db 279 GNYTCVAANKLGVTNASLVLPFGSVRGING-SISLAVPL-----WL 319

RESULT 14

US-09-135-080-2

; Sequence 2, Application US/09135080

; Patent No. 6423827

;; GENERAL INFORMATION:

;; APPLICANT: Lovitt, Pat R.

;; APPLICANT: Pimenta, Aurea

;; APPLICANT: Fischer, Itzhak

;; APPLICANT: Zhukareva, Victoria

;; TITLE OF INVENTION: Limbic System-Associated Membrane

;; TITLE OF INVENTION: Protein and DNA

;; NUMBER OF SEQUENCES: 29

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: Dechert Price & Rhoads

;; STREET: 997 Lenox Drive, Building 3, Suite 210

;; CITY: Lawrenceville

;; STATE: NJ

;; COUNTRY: USA

;; ZIP: 08543

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Diskette

;; COMPUTER: IBM Compatible

;; OPERATING SYSTEM: DOS

;; SOFTWARE: FastSEQ for Windows Version 2.0

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/09/135,080

;; FILING DATE: 17-AUG-1998

;; CLASSIFICATION: 424

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: 08/414,657

;; FILING DATE: 31-MAR-1995

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Bloom, Allen

;; REGISTRATION NUMBER: 29,135

;; REFERENCE/DOCKET NUMBER: 317743-102A

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: 609-620-3214

;; TELEFAX: 609-620-3259

;; TELEX:

;; INFORMATION FOR SEQ ID NO: 2:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 325 amino acids

;; TYPE: amino acid

;; STRANDEDNESS: single

;; TOPOLOGY: linear

;; MOLECULE TYPE: protein

;; FRAGMENT TYPE: internal

US-09-135-080-2

Query Match 5.6%; Score 80.5; DB 4; Length 325;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

CM protein - protein search, using sw mode:

Run on: September 26, 2003, 17:23:47 ; Search time 27 Seconds
(without alignments)
1507.467 Million cell updates/sec

Title: US-09-852-100A-2

Perfect score: 1439

Sequence: 1 MHILKGNPNVPRAGQKNT.....LHRLSLITNE:FRKTQLYF 269

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters: 566894

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA: *
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep: *
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep: *
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep: *
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep: *
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep: *
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep: *
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep: *
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep: *
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep: *
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep: *
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep: *
12: /cgn2_6/ptodata/2/pubpaa/US09A_NEW_PUB.pep: *
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep: *
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep: *
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep: *
16: /cgn2_6/ptodata/2/pubpaa/US10A_NEW_PUB.pep: *
17: /cgn2_6/ptodata/2/pubpaa/US10B_NEW_PUB.pep: *
18: /cgn2_6/ptodata/2/pubpaa/US10C_NEW_PUBCOMB.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1439	100.0	269	9	US-09-852-100A-2	Sequence 2, Appli
2	1439	100.0	269	10	US-09-833-503A-2	Sequence 2, Appli
3	1439	100.0	269	15	US-10-199-881-4	Sequence 2, Appli
4	201	14.0	221	11	US-09-974-879-230	Sequence 230, App
5	201	14.0	222	11	US-09-305-736-230	Sequence 230, App
6	200	13.9	221	10	US-09-833-503A-6	Sequence 6, Appli
7	200	13.9	221	11	US-09-992-600A-82	Sequence 82, Appli
8	200	13.9	221	11	US-09-924-340-82	Sequence 82, App
9	200	13.9	221	12	US-09-992-095B-82	Sequence 82, App
10	200	13.9	221	12	US-10-154-678-82	Sequence 82, App
11	200	13.9	221	12	US-09-999-570-82	Sequence 82, Appl
12	200	13.9	221	15	US-10-000-489-82	Sequence 82, Appl
13	200	13.9	221	15	US-10-000-986-82	Sequence 82, Appl
14	200	13.9	221	15	US-10-199-881-6	Sequence 6, Appli
15	200	13.9	247	11	US-09-796-753-48	Sequence 48, Appl

16	196	13.6	230	11	US-09-796-753-56	Sequence 50, Appli
17	136.5	9.5	214	10	US-09-833-503A-4	Sequence 4, Appli
18	136.5	9.5	214	15	US-10-199-881-4	Sequence 4, Appli
19	136.5	9.5	224	9	US-09-925-301-1018	Sequence 1018, Af
20	121	8.4	172	11	US-09-798-889-58	Sequence 58, Appli
21	115.5	8.0	213	11	US-09-866-050A-631	Sequence 631, App
22	94.5	6.6	166	15	US-10-156-761-10938	Sequence 10938, A
23	92.5	6.4	87	15	US-10-156-761-10936	Sequence 10936, A
24	90	6.3	443	15	US-10-156-761-9371	Sequence 9371, Ap
25	87	6.0	476	15	US-10-156-761-8693	Sequence 8693, Ap
26	85	5.9	338	10	US-09-808-602-72	Sequence 72, Appli
27	85	5.9	338	11	US-09-800-198-61	Sequence 61, Appli
28	84.5	5.9	389	9	US-09-925-297-489	Sequence 489, App
29	83.5	5.8	179	15	US-10-156-761-13418	Sequence 13418, A
30	83	5.8	764	15	US-10-177-293-27	Sequence 27, Appli
31	82.5	5.7	338	10	US-09-808-602-69	Sequence 69, Appli
32	82.5	5.7	338	11	US-09-800-198-58	Sequence 58, Appli
33	82.5	5.7	338	15	US-10-161-572-58	Sequence 58, Appli
34	82	5.7	416	15	US-10-156-761-10376	Sequence 10376, A
35	81	5.6	717	15	US-10-054-044A-2	Sequence 2, Appli
36	81	5.6	797	9	US-09-795-232-2	Sequence 2, Appli
37	80.5	5.6	373	11	US-09-284-320-17	Sequence 17, Appli
38	80	5.6	328	11	US-09-961-403-2	Sequence 2, Appli
39	80	5.6	328	14	US-10-011-859-28	Sequence 28, Appli
40	80	5.6	476	15	US-10-080-917-7	Sequence 7, Appli
41	80	5.6	660	15	US-10-294-055-8	Sequence 8, Appli
42	79.5	5.5	987	15	US-10-156-761-4515	Sequence 14515, A
43	79	5.5	797	12	US-10-151-813-124	Sequence 124, App
44	78.5	5.4	245	11	US-09-800-198-92	Sequence 92, Appli
45	78	5.4	304	10	US-09-738-626-6200	Sequence 6200, Ap

ALIGNMENTS

RESULT 1

US-09-852-100A-2
; Sequence 2, Application US/09852100A
; Patent No. US20020058267A1
; GENERAL INFORMATION:
; APPLICANT: American Home Products
; TITLE OF INVENTION: Beta-amyloid Peptide-Binding Proteins and Polynucleotides En
; FILE REFERENCE: AHP981261P2
; CURRENT FILING DATE: 2001-05-09
; PRIOR FILING DATE: 1998-10-14
; PRIOR APPLICATION NUMBER: US 09/172,990
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: US 60/104,104
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: PTC/US99/21621
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: US 09/060,609
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: US 60/064,583
; PRIOR FILING DATE: 1997-04-16
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-852-100A-2

Query Match 100.0%; Score 1439; DB 9; Length 269;

Best Local Similarity 100.0%; Pred. No. 2e-135;

Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHILKGNPNVPRAGQKNTRDGTLGYPMRGPFKNLALLPFLGSGSGSEKYSV 60

DB 1 MHILKGNPNVPRAGQKNTRDGTLGYPMRGPFKNLALLPFLGSGSGSEKYSV 60

QY 61 SKMAAWPSPGSAPEAVTARLVGVSVITTPMGAVATSAGGESLKCEDLKVGQYIC 120

```
Db 61 SKMAAAMPSPAPAVTARLVGLVLFVSVITGPGAVATSAGGEESLKCEDLVGVQYIC 120
Qy 121 KDPKINDATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFPPISCRNVNG 180
Db 121 KDPKINDATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFPPISCRNVNG 180
Qy 181 YSKYKVAVALSLFLGWLGDREYLVGYPALGLLKFCFTVGFVCGIGSLDFILISMQIVGPSDG 240
Db 181 YSKYKVAVALSLFLGWLGDREYLVGYPALGLLKFCFTVGFVCGIGSLDFILISMQIVGPSDG 240
Qy 241 SSIYIDYGTTRLRLSITNETFRKTQLYP 269
Db 241 SSIYIDYGTTRLRLSITNETFRKTQLYP 269

RESULT 2
US-09-833-503A-2
: Sequence 2, Application US/09833503A
: Patent No. US20020146760A1
: GENERAL INFORMATION:
: APPLICANT: Ozenberger, Bradley A
: APPLICANT: Kajkowski, Eileen M
: APPLICANT: Lo, Chang-Hsiung F
: APPLICANT: American Home Products Corporation
: TITLE OF INVENTION: Polynucleotides Encoded By Them, and Methods of Using
: TITLE OF INVENTION: Polynucleotides Encoded By Them, and Methods of Using
: TITLE OF INVENTION: Same
: FILE REFERENCE: AHP98165-00PCT
: CURRENT APPLICATION NUMBER: US/09/833,503A
: CURRENT FILING DATE: 2000-10-13
: PRIOR APPLICATION NUMBER: 50/104,104
: PRIOR FILING DATE: 1998-10-13
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 2
: LENGTH: 269
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-833-503A-2

Query Match 100.0%; Score 1439; DB 10; Length 269;
Best Local Similarity 100.0%; Pred. No. 2e-135;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHILKSPNVIIPRAHQKNTRRDGTGLYPMRGPFKNLALPFLSLPLGGGGSGSGEKVSU 60
Db 1 MHILKSPNVIIPRAHQKNTRRDGTGLYPMRGPFKNLALPFLSLPLGGGGSGSGEKVSU 60
Qy 61 SKMAAAMPSPAPAVTARLVGLVLFVSVITGPGAVATSAGGEESLKCEDLVGVQYIC 120
Db 61 SKMAAAMPSPAPAVTARLVGLVLFVSVITGPGAVATSAGGEESLKCEDLVGVQYIC 120
Qy 121 KDPKINDATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFPPISCRNVNG 180
Db 121 KDPKINDATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFPPISCRNVNG 180
Qy 181 YSKYKVAVALSLFLGWLGDREYLVGYPALGLLKFCFTVGFVCGIGSLDFILISMQIVGPSDG 240
Db 181 YSKYKVAVALSLFLGWLGDREYLVGYPALGLLKFCFTVGFVCGIGSLDFILISMQIVGPSDG 240
Qy 241 SSIYIDYGTTRLRLSITNETFRKTQLYP 269
Db 241 SSIYIDYGTTRLRLSITNETFRKTQLYP 269

RESULT 3
US-10-199-881-2
: Sequence 2, Application US/10199881
: Publication No. US20030096356A1
: GENERAL INFORMATION:
: APPLICANT: Wyeth
: TITLE OF INVENTION: No. US20030096356A1 G-Protein-Coupled Receptor-Like Proteins and
```

```
: TITLE OF INVENTION: Encoded by Them, and Methods of Using Same"
: FILE REFERENCE: AHP98165C1
: CURRENT APPLICATION NUMBER: US/10/199,881
: CURRENT FILING DATE: 2002-07-18
: PRIOR APPLICATION NUMBER: PCT/US99/21621
: PRIOR FILING DATE: 1999-10-13
: PRIOR APPLICATION NUMBER: US 90/833,5081
: PRIOR FILING DATE: 2001-12-04
: PRIOR APPLICATION NUMBER: US 60/104,104
: PRIOR FILING DATE: 1998-10-13
: NUMBER OF SEQ ID NOS: 45
: SOFTWARE: Patent In version 3.1
: SEQ ID NO 2
: LENGTH: 269
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-199-881-2

Query Match 100.0%; Score 1439; DB 15; Length 269;
Best Local Similarity 100.0%; Pred. No. 2e-135;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHILKSPNVIIPRAHQKNTRRDGTGLYPMRGPFKNLALPFLSLPLGGGGSGSGEKVSU 60
Db 1 MHILKSPNVIIPRAHQKNTRRDGTGLYPMRGPFKNLALPFLSLPLGGGGSGSGEKVSU 60
Qy 61 SKMAAAMPSPAPAVTARLVGLVLFVSVITGPGAVATSAGGEESLKCEDLVGVQYIC 120
Db 61 SKMAAAMPSPAPAVTARLVGLVLFVSVITGPGAVATSAGGEESLKCEDLVGVQYIC 120
Qy 121 KDPKINDATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFPPISCRNVNG 180
Db 121 KDPKINDATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFPPISCRNVNG 180
Qy 181 YSKYKVAVALSLFLGWLGDREYLVGYPALGLLKFCFTVGFVCGIGSLDFILISMQIVGPSDG 240
Db 181 YSKYKVAVALSLFLGWLGDREYLVGYPALGLLKFCFTVGFVCGIGSLDFILISMQIVGPSDG 240
Qy 241 SSIYIDYGTTRLRLSITNETFRKTQLYP 269
Db 241 SSIYIDYGTTRLRLSITNETFRKTQLYP 269

RESULT 4
US-09-974-879-230
: Sequence 230, Application US/09974879
: Publication No. US20030028003A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: 125 Human Secreted Proteins
: FILE REFERENCE: P2020P2
: CURRENT APPLICATION NUMBER: US/09/974,879
: CURRENT FILING DATE: 2001-10-12
: PRIOR APPLICATION NUMBER: US 60/239,893
: PRIOR FILING DATE: 2000-10-13
: PRIOR APPLICATION NUMBER: US 09/818,683
: PRIOR FILING DATE: 2001-03-28
: PRIOR APPLICATION NUMBER: US 09/305,736
: PRIOR FILING DATE: 1999-05-05
: PRIOR APPLICATION NUMBER: PCT/US98/23435
: PRIOR FILING DATE: 1998-11-04
: PRIOR APPLICATION NUMBER: US 60/064,911
: PRIOR FILING DATE: 1997-11-07
: PRIOR APPLICATION NUMBER: US 60/064,912
: PRIOR FILING DATE: 1997-11-07
: PRIOR APPLICATION NUMBER: US 60/064,983
: PRIOR FILING DATE: 1997-11-07
: PRIOR APPLICATION NUMBER: US 60/064,900
: PRIOR FILING DATE: 1997-11-07
: PRIOR APPLICATION NUMBER: US 60/064,988
: PRIOR FILING DATE: 1997-11-07
: PRIOR APPLICATION NUMBER: US 60/064,987
: PRIOR FILING DATE: 1997-11-07
```

```

Query Match      13.9%: Score 200; DB 10; Length 221;
Best Local Similarity 45.7%; Pred. No. 7.4e-12;
Matches 53; Conservative 12; Mismatches 39; Indels 12; Gaps 5;

QY 135 CTNYTA--HVSC----FPAPNITCKDSGNETHFTTGNVEGFFPISCRNVNNGYSYKVAVA 183
      ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 112 CTNSTSCMTVSPCRYPFA-NCTVRD-----HVHCLGNRT-FPKMLCYNNTGGYKWKSTALA 163

QY 189 LSLFLGWLGAORFYLGYLPALGLLNFCEVGVGCGIGSLDFILISMQIVGSPDGSSEYI 244
      ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 166 LSLITLGGFCADRFLYGQWRREGSLKLFSGGLGIWTLIDVLLIGVYGVGADGSLEYI 221

```



```
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 182.US1.REG
; CURRENT APPLICATION NUMBER: US/10/54,578
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/296,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 82
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: 1..32
; US-09-999-570-82
```

```
Query Match 13.9%; Score 200; DB 12; Length 221;
Best Local Similarity 45.7%; Pred. No. 7.4e-12;
Matches 53; Conservative 12; Mismatches 39; Indels 12; Gaps 5;

QY 135 CTNYTA--HVSC-----FPAPNTCKDSSGNETHFTGNEVGFPKPISCRNVNGYSYKVAVA 188
||| | :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 112 CTNSTSMTVSCPRQRYPA-NCTVRD----HVHCLGNRT-FPKMLYCNWTTGGYKWSIALA 165
||| | :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 189 LSLFLGWLGAADRFYLGYPALGLLKFTVCGFCGIGSLIDFILISMQIVGPGDSSSYI 244
||| | :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 166 LSITLGGFADRFYLGQWREGLGKLFSGGLGIWTLIDVLLIGVGVGPADGSLYI 221
||| | :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
```

```
RESULT 11
US-09-999-570-82
; Sequence 82, Application US/09999570
; Publication No. US20030170628A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: G-091US080DIV
; CURRENT APPLICATION NUMBER: US/09/999,570
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 82
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: 1..32
; US-09-999-570-82
```

Query Match 13.9%; Score 200; DB 12; Length 221;

```
Best Local Similarity 45.7%; Pred. No. 7.4e-12;
Matches 53; Conservative 12; Mismatches 39; Indels 12; Gaps 5;

QY 135 CTNYTA--HVSC-----FPAPNTCKDSSGNETHFTGNEVGFPKPISCRNVNGYSYKVAVA 188
||| | :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 112 CTNSTSMTVSCPRQRYPA-NCTVRD----HVHCLGNRT-FPKMLYCNWTTGGYKWSIALA 165
||| | :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 189 LSLFLGWLGAADRFYLGYPALGLLKFTVCGFCGIGSLIDFILISMQIVGPGDSSSYI 244
||| | :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 166 LSITLGGFADRFYLGQWREGLGKLFSGGLGIWTLIDVLLIGVGVGPADGSLYI 221
||| | :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
```

```
RESULT 12
US-10-000-489-82
; Sequence 82, Application US/10000489
; Publication No. US2003009201A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US6.DIV
; CURRENT APPLICATION NUMBER: US/10/000,489
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 82
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: 1..32
; US-10-000-489-82
```

```
Query Match 13.9%; Score 200; DB 15; Length 221;
Best Local Similarity 45.7%; Pred. No. 7.4e-12;
Matches 53; Conservative 12; Mismatches 39; Indels 12; Gaps 5;

QY 135 CTNYTA--HVSC-----FPAPNTCKDSSGNETHFTGNEVGFPKPISCRNVNGYSYKVAVA 186
||| | :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 112 CTNSTSMTVSCPRQRYPA-NCTVRD----HVHCLGNRT-FPKMLYCNWTTGGYKWSIALA 165
||| | :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 189 LSLFLGWLGAADRFYLGYPALGLLKFTVCGFCGIGSLIDFILISMQIVGPGDSSSYI 244
||| | :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 166 LSITLGGFADRFYLGQWREGLGKLFSGGLGIWTLIDVLLIGVGVGPADGSLYI 221
||| | :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
```

```
RESULT 13
US-10-000-985-82
; Sequence 82, Application US/10000985
; Publication No. US20030096247A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US9.DIV
; CURRENT APPLICATION NUMBER: US/10/000,986
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
```


Db 138 CTNSTSCMTVSCPQRYPANCTVRD----HVHCLGNRT-EPKMLYCNWTTGGYKWSIALA 191
Qy 189 LSLFLGWLGCADREYLGYPALGJLKFCVGFCCGIGSLIDFELLSMQLVGFSUSSYI 244
Db 192 LSTITGGFGADREYLGQWRGSLGKLFSPGGLGIWTLIDVLLIGVGYVGPADGSLYI 247

Search completed: September 26, 2003, 17:28:52
Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2003, 17:22:11 : Search time 43 seconds
(without alignments)
691.613 Million cell updates/sec

Title: US-09-852-100A-2

Perfect score: 1439

Sequence: 1 MHILKSPNVLPAHCOXNF.....TLRLSLTINTEFRKTLQLY: 263

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_76: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	167.5	11.6	573	2 S44605	C02F5.3 protein -
2	159.5	11.1	753	2 T28787	hypothetical prote
3	95	6.6	309	2 H75286	hypothetical prote
4	92	6.4	456	2 S55681	hypothetical prote
5	91	6.3	1324	1 VGIH59	E2 glycoprotein: fr
6	89.5	6.2	690	2 T08694	hypothetical prote
7	87.5	6.1	547	2 H75632	Na(+)-linked D-ala
8	87.5	6.1	1207	2 T23754	hypothetical prote
9	85.5	5.9	348	1 S20911	alcohol dehydrogen
10	84	5.8	446	2 T35005	probable integral
11	83.5	5.8	487	2 C70574	probable atrop2 pro
12	83	5.8	137	2 H75447	hypothetical prote
13	83	5.8	348	1 S32521	alcohol dehydrogen
14	83	5.8	764	1 BBHG	component factor
15	83	5.8	1274	2 T10729	transferrin-like p
16	82.5	5.7	338	2 JC4776	limbic-system asso
17	82	5.7	409	2 T03788	laccase (EC 1.10.3
18	81.5	5.7	504	2 T37184	probable monooxyge
19	81.5	5.7	1266	2 AC2695	conserved hypothet
20	81.5	5.7	1266	2 G97477	hypothetical prote
21	81	5.6	443	2 G90388	thermopline protei
22	81	5.6	513	2 A81850	probable integral
23	81	5.6	513	2 C81092	oxalate/formate an
24	81	5.6	758	2 D87369	beta-D-glucosidase
25	81	5.6	1176	2 T49482	hypothetical prote
26	80.5	5.6	134	2 D83595	hypothetical prote
27	80.5	5.6	223	2 F72803	qp32 protein - Myc
28	80.5	5.6	1237	2 A54080	protein-tyrosine-p
29	80	5.6	171	2 T01466	iron-su-fur cofact

ALIGNMENTS

RESULT 1

S44605

C02F5.3 protein - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 02-Feb-2001

C:Accession: S44605

R:Anderson, K.

Submitted to the EMBL Data Library, May 1993

A:Description: Sequence of the C. elegans cosmid C02F5.

A:Reference number: S44603

A:Accession: S44605

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-573 <AND>

A:Cross-references: EMBL:L14745; NID:g289607; PID:g289610

C:Genetics:

A:Introns: 224/2; 304/1; 363/3; 390/3; 503/2

C:Superfamily: translation elongation factor Tu homology

C:Keywords: GTP binding; nucleotide binding; P-loop

F:68-183/Domain: translation elongation factor Tu homology <F50>

F:69-76/Region: nucleotide-binding motif A (P-loop)

F:246-249/Region: GTP-binding NKXD motif

Query Match 11.6%; Score 167.5; DB 2; Length 573;

Best Local Similarity 27.9%; Pred No. 1.3e-06;

Matches 50; Conservative 23; Mismatches 57; Indels 49; Gaps 5;

QY 9C VTTGPGAVATSAAGSESLKEDLVKGVQYICKDP-----K 124

DB 41S VSTNPLGPV-----VECRFLENSFLCEDPVLPGQGTGQOPANESFRNEGKCLK 465

QY 22S :NDATQEPVNCNTYAHVSCFPAPNITCKDSNGNETHGTGNEVGFFKPISCRNNVNGYSYK 184

DB 466 MGYRAEDVEFTN-----VACRVLPCTEC---HGPRP-----FTRTPTCIYNGHYEL 510

QY 18S VAVALSLEFLGADRFYLGYPALGLKFCVTCGICGLIDFILSMQIVGPSDGSY 243

DB 511 TLLYSIFLGVAVDREFLGYSAMAVGLMTLGGFGIWMIVDIFLLVGLVGPADSSM 569

RESULT 2

T28787

hypothetical protein C4D11.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T28787

R:Gattung, S.; Magg1, L.

Submitted to the EMBL Data Library, May 1997

A:Description: The sequence of C. elegans cosmid C4D11.

A:Reference number: Z20522

A:Accession: T28787

A>Status: preliminary; translated from GB/EMBL/DDBJ

```

Query Match      6.3%   Score 91; DB 1; Length 1324;
Best Local Similarity 23.6%; Pred. No. 13;
Matches 59; Conservative 27; Mismatches 94; Indels 70; Gaps 13;

QY    25 TGLYPWRG-PKKNLALLP-----FSLPLGGGSGCEKVSVSKMAAHPGPSA-- 73
       ||||| : | : |||||         | | | | | : | : ||| : |
Db     66 TCYYPVDSCKFRNLALTGTNSVLSWFOPLYLNQFDNGIFAK--VNLIKSTPSGATAYF 123
       ||||| : | : |||||         | | | | | : | : ||| : |

QY    74 PEAVTARLVGVLFVSVTTPGMGAVALTSAGSEESLKCDLKVGY-ICKDKPKINDATQEP 132
       ||||| : | : |||||         | | | | | : | : ||| : |
Db    124 PRIVIGSLFGVTSY-TVVIEPYNGVIMAS-----VCQYTICQLP----- 161

```

A:Title: Primary structure of the glycoprotein E2 of coronavirus HKV-A39 and identification of the E2 glycoprotein of coronavirus HKV-A39 and identification of the E2 glycoprotein of coronavirus HKV-A39

A:Accession number: A27402; M01D:86072086; PMID:2825419

A:Accession: A27402

A:Molecule type: genomic RNA

A:Residues: 1-1324 <L0Y>

C:Superfamily: coronavirus E2 glycoprotein

C:Keywords: glycoprotein; transmembrane protein

F:1-16/Domain: signal sequence #status predicted <SIG>

F:17-1324/Product: E2 glycoprotein #status predicted <E2G>

F:17-717/Product: 90B glycoprotein #status predicted <E2G>

F:718-1324/Product: 90A glycoprotein #status predicted <E2G>

F:1266-1286/Product: transmembrane #status predicted <E2G>

F:31-60-192/247-357,435-442-530-625,657,665-688,737,754,784,844-893,1126-1180,1190-1209,1219-1230-1231-1232-1233-1234-1235-1236-1237-1238-1239-1240-1241-1242-1243-1244-1245-1246-1247-1248-1249-1250-1251-1252-1253-1254-1255-1256-1257-1258-1259-1260-1261-1262-1263-1264-1265-1266-1267-1268-1269-1270-1271-1272-1273-1274-1275-1276-1277-1278-1279-1280-1281-1282-1283-1284-1285-1286-1287-1288-1289-1290-1291-1292-1293-1294-1295-1296-1297-1298-1299-1300-1301-1302-1303-1304-1305-1306-1307-1308-1309-1310-1311-1312-1313-1314-1315-1316-1317-1318-1319-1320-1321-1322-1323-1324-1325-1326-1327-1328-1329-1330-1331-1332-1333-1334-1335-1336-1337-1338-1339-1340-1341-1342-1343-1344-1345-1346-1347-1348-1349-1350-1351-1352-1353-1354-1355-1356-1357-1358-1359-1360-1361-1362-1363-1364-1365-1366-1367-1368-1369-1370-1371-1372-1373-1374-1375-1376-1377-1378-1379-1380-1381-1382-1383-1384-1385-1386-1387-1388-1389-1390-1391-1392-1393-1394-1395-1396-1397-1398-1399-1400-1401-1402-1403-1404-1405-1406-1407-1408-1409-1410-1411-1412-1413-1414-1415-1416-1417-1418-1419-1420-1421-1422-1423-1424-1425-1426-1427-1428-1429-1430-1431-1432-1433-1434-1435-1436-1437-1438-1439-1440-1441-1442-1443-1444-1445-1446-1447-1448-1449-1450-1451-1452-1453-1454-1455-1456-1457-1458-1459-1460-1461-1462-1463-1464-1465-1466-1467-1468-1469-1470-1471-1472-1473-1474-1475-1476-1477-1478-1479-1480-1481-1482-1483-1484-1485-1486-1487-1488-1489-1490-1491-1492-1493-1494-1495-1496-1497-1498-1499-1500-1501-1502-1503-1504-1505-1506-1507-1508-1509-1510-1511-1512-1513-1514-1515-1516-1517-1518-1519-1520-1521-1522-1523-1524-1525-1526-1527-1528-1529-1530-1531-1532-1533-1534-1535-1536-1537-1538-1539-1540-1541-1542-1543-1544-1545-1546-1547-1548-1549-1550-1551-1552-1553-1554-1555-1556-1557-1558-1559-1560-1561-1562-1563-1564-1565-1566-1567-1568-1569-1570-1571-1572-1573-1574-1575-1576-1577-1578-1579-1580-1581-1582-1583-1584-1585-1586-1587-1588-1589-1590-1591-1592-1593-1594-1595-1596-1597-1598-1599-1600-1601-1602-1603-1604-1605-1606-1607-1608-1609-1610-1611-1612-1613-1614-1615-1616-1617-1618-1619-1620-1621-1622-1623-1624-1625-1626-1627-1628-1629-1630-1631-1632-1633-1634-1635-1636-1637-1638-1639-1640-1641-1642-1643-1644-1645-1646-1647-1648-1649-1650-1651-1652-1653-1654-1655-1656-1657-1658-1659-1660-1661-1662-1663-1664-1665-1666-1667-1668-1669-1670-1671-1672-1673-1674-1675-1676-1677-1678-1679-1680-1681-1682-1683-1684-1685-1686-1687-1688-1689-1690-1691-1692-1693-1694-1695-1696-1697-1698-1699-1700-1701-1702-1703-1704-1705-1706-1707-1708-1709-1710-1711-1712-1713-1714-1715-1716-1717-1718-1719-1720-1721-1722-1723-1724-1725-1726-1727-1728-1729-1730-1731-1732-1733-1734-1735-1736-1737-1738-1739-1740-1741-1742-1743-1744-1745-1746-1747-1748-1749-1750-1751-1752-1753-1754-1755-1756-1757-1758-1759-1760-1761-1762-1763-1764-1765-1766-1767-1768-1769-1770-1771-1772-1773-1774-1775-1776-1777-1778-1779-1780-1781-1782-1783-1784-1785-1786-1787-1788-1789-1790-1791-1792-1793-1794-1795-1796-1797-1798-1799-1800-1801-1802-1803-1804-1805-1806-1807-1808-1809-1810-1811-1812-1813-1814-1815-1816-1817-1818-1819-1820-1821-1822-1823-1824-1825-1826-1827-1828-1829-1830-1831-1832-1833-1834-1835-1836-1837-1838-1839-1840-1841-1842-1843-1844-1845-1846-1847-1848-1849-1850-1851-1852-1853-1854-1855-1856-1857-1858-1859-1860-1861-1862-1863-1864-1865-1866-1867-1868-1869-1870-1871-1872-1873-1874-1875-1876-1877-1878-1879-1880-1881-1882-1883-1884-1885-1886-1887-1888-1889-1890-1891-1892-1893-1894-1895-1896-1897-1898-1899-1900-1901-1902-1903-1904-1905-1906-1907-1908-1909-1910-1911-1912-1913-1914-1915-1916-1917-1918-1919-1920-1921-1922-1923-1924-1925-1926-1927-1928-1929-1930-1931-1932-1933-1934-1935-1936-1937-1938-1939-1940-1941-1942-1943-1944-1945-1946-1947-1948-1949-1950-1951-1952-1953-1954-1955-1956-1957-1958-1959-1960-1961-1962-1963-1964-196

QY 133 VNCINTYATVSCFPAPNITCKDSSGNETHFTGNE-VQFF-----KPLSCRNNGYGYKVVAV 187
 Db 162 -----YTDCKNTN-----GNKLIGFHWFDVKKPICVLKRNETLVNA 199
 QY 188 ALSFLGWLQADRYLYG---PALGLLKFTVGCIGSLIDFLILSMQIVGPSDSSSYI 244
 Db 200 DAFYFHYQHGSGTFYAYADKPSATIFLSVY----ICDILTYVYVLPFCINPTAGSTFA 255
 QY 245 IDYVGIRLTR 254
 Db 256 PRYWTPLVK 265
 RESULT 6
 T08604
 hypothetical protein GRR1 - soybean
 C:Species: Glycine max (soybean)
 C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999
 C:Accession: T08604
 R:Chen, W.; Atherly, A.
 submitted to the EMBL Data Library, August 1997
 A:Reference number: Z15438
 A:Accession: T08604
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-690 <CH>
 A:Cross-references: EMBL:AF019910; NID:g2407789; PID:g2407790
 A:Experimental source: variety L85-3044; root
 C:Genetics:
 A:Gene: grr1

Query Match 6.2%; Score 89.5; DB 2; Length 690;
 Best Local Similarity 24.6%; Pred. No. 8.4;
 Matches 43; Conservative 22; Mismatches 59; Indels 51; Gaps 7;

QY 85 LWFVSITGPMGAVATSAGGSESLKCEDEKYGQYICKDPKINDATQEPV--NCTNYTAHV 142
 Db 233 LWQVA-TVGDVGLTEIASGCHQLEKLD-----LCKCPNISDKTLIAVAKNCFN-LAE 283
 QY 143 SCFPAPNITCKDSSGNETHFTGNEVGFKPKISCRNVNGYKVAVALSLFLCWLQADRFY 202
 Db 284 SIQCPNI-----GNEQLQAIKCPNLRISIKNCSGVQ----- 319
 QY 203 LGYPALGLLKFTVGCIGSLIDFL--ISMQIVGPSDSSSYI:DNVGTPLTR: 255
 Db 320 -----GVAGLLSSASFALTKVKLESITVSDLSAVZGHYGVAVTCL 360

RESULT 7

H75632
 Na(+)-linked D-alanine glycine permease - Deinococcus radiodurans (strain K1)
 C:Species: Deinococcus radiodurans
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
 C:Accession: H75632
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, L.; Zalewski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 285, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567255
 A:Accession: H75632
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-547 <WH>
 A:Cross-references: GB:AF001826; NID:g6460827; PIDN:AAFL2563.1; PID:g6460859; TIGR:DRB01
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DRB0133
 A:Map position: megapiasmid
 A:Genome: plasmid
 A:Note: plasmid MPI
 C:Superfamily: sodium-dependent D-alanine/glycine transport protein.

Query Match 6.1%; Score 87.5; DB 2; Length 547;
 Best Local Similarity 25.6%; Pred. No. 9.6;
 Matches 57; Conservative 19; Mismatches 52; Indels 95; Gaps 13;

QY 43 SLPLGGGGSGSGEKYSVSKMAAA-----WPS--GFSAPAEAVZARLVGV----- 84
 Db 29 SRPLSSESGSSAQEPWGRPLPAALVFTGLLGAVSASQAQSPVDERINAVVTPVSHFLS 88
 QY 85 -LWFVSVTGTP-----WGAVATSGGSESLKCE-----DLKVGQYIC 120
 Db 89 GLIFASISVGEAQVPLIVVWLAVA-----AVCTLSFREFVNIWGFKHGIDLVRGY-- 139
 QY 121 KDPKINDATQEPVNCTNYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFKPKISCRNVNG 180
 Db 140 ----GNDA-----DAGQWTHFVALTAVNSGTVGLGNIG 170
 QY 181 YSKVAVALSLFLGWLQADRFYLGYPALGLL----KF--CTVG 217
 Db 171 ---VAVALS--LGGPGATFMMI---LVGLLSMSTKFEVCTLG 204
 RESULT 8
 T23754
 hypothetical protein T05C12.10 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T23754; T24513
 R:Thomas, K.
 submitted to the EMBL Data Library, June 1995
 A:Reference number: Z19793
 A:Accession: T23754
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1207 <ML>
 A:Cross-references: EMBL:Z49968; PIDN:CAA90265.1; GSPDB:GN00020; CESP:T05C12.10
 A:Experimental source: clone M110
 R:Burton, J.
 submitted to the EMBL Data Library, October 1995
 A:Reference number: Z19901
 A:Accession: T24513
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1207 <M12>
 A:Cross-references: EMBL:Z66500; PIDN:CAA91313.1; GSPDB:GN00020; CESP:T05C12.10
 A:Experimental source: clone T05C12
 C:Genetics:
 A:Gene: CESP:T05C12.10
 A:Map position: 2
 A:Mutrons: 31/3; 87/2; 141/3; 160/2; 203/3; 267/1; 776/2; 794/2; 834/2; 1086/1; 1143

Query Match 6.1%; Score 87.5; DB 2; Length 1207;
 Best Local Similarity 21.1%; Pred. No. 23;
 Matches 47; Conservative 26; Mismatches 83; Indels 67; Gaps 8;

QY 2 HILKSPNPIVPAHGOKNTRDGTGLYPMRGPKFNALLPFLPLGGGGSGSGEKYSVS 61
 Db 848 HNAESSASGIPLYQARSNTYNGAPVPAPGS-----GATSGTSGSGTSESVT 894
 QY 62 KMAANWSPGS-----APEAVTARLVGV:MFVSVITGPMGAVATSAG 100
 Db 897 NGSGATESGSGTSGTGTGTSGTSGTSSAARTSSIAGDAQAAVLADTPGAAGAAG 954
 QY 104 GE-----ESL-----KCEDLKVGQYICKDPKINDAT-----QEPVNCNNTYA 140
 Db 957 GRNCFNSADSLVITVTGQKRMDLQIGDYVLVPSGGNVLKYKVEFYHREPKTRNF-- 1054
 QY 141 HVSCFPAPNITCKDSSGNETHFTGNEVGFKPKIS--CRNVNGYS 182
 Db 1025 -----VLYTKSGRKLTLTGRLH---LPVAECQSVEQYT 1045

RESULT 9
 S20911

RESULT 10
T35005
probable integral membrane transporter - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 17-Mar-2000
C:Accession: T35005
R:Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, December 1998
A:Reference number: Z21564
A:Accession: T35005
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-446 <SEE>
A:Cross-references: EMBL:AL034443; PTDN:CAA22367.1; GSFPB:GN00070; SCOEDB:SC4B5.13
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SC4B5.13
C:Superfamily: hypothetical protein c0i03

Query Match 5.88; Score 84; DB 2: Length 446;
Best Local Similarity 19.7%; Pred. No. 15;
Matches 52; Conservative 34; Mismatches 110; Indels 68; Gaps 9;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2003, 17:15:17 : Search time 23 Seconds
(without alignments)
550.008 Million cell updates/sec

Title: US-09-852-100A-2

Perfect score: 1439

Sequence: 1 MHILKSPNPVIRAHGKQNT.....IRLRLSLTNEIRKTKQLYP 259

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127363

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	167.5	11.6	573	1 YKK3_CAEEL	P34280 caenorhabdi
2	90	6.3	1324	1 VGL2_CVNA5	P11224 murine corb
3	89.5	6.2	515	1 EF15_PORPU	P50257 porphyra pu
4	86.5	6.0	151	1 LCT2_MOUSE	O88803 mus musculu
5	85.5	5.9	348	1 ADH2_KLUJA	P49383 kluyveromyc
6	85	5.9	338	1 LAMP_RAT	Q62813 rattus norv
7	83.5	5.8	487	1 Y346_MYCTU	O06297 mycobacteri
8	83	5.8	349	1 ADH1_KLUJA	Q67289 kluyveromyc
9	83	5.8	764	1 CPAB_HUMAN	P00751 homo sapien
10	82.5	5.7	338	1 LAMP_HUMAN	O13445 homo sapien
11	80.5	5.6	223	1 VG32_BPMO2	O64226 mycobacteri
12	80.5	5.6	455	1 EN11_HUMAN	Q99805 homo sapien
13	80	5.6	328	1 IBP2_HUMAN	P18065 homo sapien
14	80	5.6	450	1 LIPP_PIG	P00591 sus scrofa
15	80	5.6	638	1 OARI_LYNST	O77408 lymanaca sta
16	79.5	5.5	489	1 ANSP_MYCTU	Q33261 mycobacteri
17	79.5	5.5	655	1 POLG_HRV3	Q82081 human rhino
18	79.5	5.5	910	1 AUX1_BOVIN	Q27974 bos taurus
19	79.5	5.5	3432	1 POLG_JAEV1	P27395 j domestico po
20	79.5	5.5	3432	1 POLG_JAEV5	P19110 j domestico po
21	79.5	5.5	3432	1 POLG_JAEV7	P32886 j domestico po
22	79.5	5.5	4367	1 DYHC_NEUCR	P45443 neuropept
23	79	5.5	468	1 LIP2_MOUSE	P17892 mus musculu
24	79	5.5	525	1 STP_SPIOL	Q03411 spinacia ol
25	79	5.5	797	1 SGA5_HUMAN	Q93455 homo sapien
26	78.5	5.5	245	1 CR53_HORSE	O19010 equus caball
27	78.5	5.5	398	1 MUB1_XENLA	P38565 xenopus lae
28	78.5	5.5	5376	1 ZAN_MOUSE	O88799 mus musculu
29	78	5.4	557	1 YP85_MYCTU	Q50636 mycobacteri
30	77.5	5.4	374	1 AKOC_METHH	O26843 methanobact
31	77.5	5.4	2179	1 POLG_HRV14	P03303 human rhino
32	77	5.4	365	1 IDI2_METMA	Q6pw37 methanosarc
33	77	5.4	1302	1 NRG_DROME	P20241 drosophila

34 76.5 5.3 772 1 PMIP_YEAST
35 76 5.3 379 1 METX_NEIMB
36 76 5.3 467 1 LIPI1_HUMAN
37 76 5.3 484 1 GATA_PSEAE
38 76 5.3 796 1 STR3_MOUSE
39 76 5.3 1556 1 VG67_HSVII
40 75.5 5.2 199 1 WRBA_YERPE
41 75.5 5.2 211 1 CLD7_MOUSE
42 75.5 5.2 350 1 ADH1_KLUJA
43 75.5 5.2 484 1 PER2_VOLUCA
44 75.5 5.2 592 1 CEA_CITFR
45 75.5 5.2 656 1 PER_DROSI

ALIGNMENTS

RESULT 1
YKK3_CAEEL
ID YKK3_CAEEL STANDARD: PKT; 573 AA.
AC P34280:1994 (Rel. 28, Created)
CT 01-FEB-1994 (Rel. 28, Last sequence update)
CT 01-FEB-1994 (Rel. 28, Last sequence update)
CT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical GTP-binding protein C02F5.3 in chromosome III.
GN C02F5.3
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID:6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A., Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A., Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L., Jones M., Kershaw J., Kirsten J., Laister N., Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M., Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R., Sims M., Smaildon N., Smith A., Smith M., Sonhammer E., Staden R., Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R., Watson A., Weinstein L., Wilkinson-Sproat J., Wohldman P.;
R "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.";
RI elegans.";
RL Nature 368:32-38(1994).
CC -!- S-MILARITY: Belongs to the GTP1 / OBG family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC FMBL; L14745; AAA27918.1;
DR PIR; S44605; S44605.
DR WormPep; C02F5.3; CE00039.
DR InterPro; IPR006074; GTP1/OBG_dom.
DR InterPro; IPR006073; GTP1_OBG.
DR InterPro; IPR006169; GTP1_OBG_sub.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF01018; GTP1_OBG.
DR Pfam; PF02824; TGS.
DR Pfam; PF05154; TM2.
DR PRINTS; PR00326; GTP1_OBG.
DR TIGRFAMs; TIGR00231; small_gtp.
DR PROSITE; PS00905; GTP1_OBG.
KW Hypothetical protein; GTP-binding.

FT	CARBOHYD	435	435	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	530	530	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	625	625	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	657	657	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	665	665	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	688	688	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	737	737	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	754	754	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	893	893	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1180	1180	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1190	1190	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1209	1209	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1225	1225	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1246	1246	N-LINKED (GLCNAC. . .) (POTENTIAL).
SEQ	SEQUENCE	1324 AA;	146019 MW;	E198EF8F0BCDBF0E CRC64;

Query Match 6.3%; Score 90; DB 1; Length 1324;
 Best Local Similarity 23.6%; Pred. No. 5.6;
 Matches 59; Conservative 27; Mismatches 94; Indels 70; Gaps 13

QY	25	TGLYPMRG-PFKNALLP-----FSLPLGGGSGSGEKYSVKMAAWPSGSA--	73
DB	66	TCGYPDGSKFKNALRGTNSVSWFPPYLNQFNQIFAK--VQMLKTIPTSGATAYF	123
QY	74	PEAVTARLVGVLFVSVTTPGWAVATSGAESLKGEDLVKGQY-ICKQPKINDATQEP	132
DB	124	PRIVIGSLFGYTSY-TVVIEPYNGVMAAS-----VCQYTTICLP-----	161
QY	133	VNCTNYTAHVSCFPAPNITCKDSSGNETHFTGNE-VGFF-----KPISCRNVNGYSYKVAV	197
DB	162	-----YTDCCKPNTN-----GNKLIGWFHFDVKPP-CVLKRNFTLNVNA	199
QY	188	ALSFLGMLGADRFVLGY---PALGLLKFCVTGFCGIGSLIDFLILSMQVIGPSDGSYY	244
DB	200	DAFYHFHYOHGGTFYAYYADKPESATITELFSVY-----IGDILTYQVYVLPFCNPTAGSIFP	255
QY	245	IDYVGTRLTR 254	
DB	256	PRYVWTPLVK 265	

RESULT 3
 EFIS_PORPU STANDARD: PRT; 515 AA.
 AC P50257;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Elongation factor 1-alpha S (EF-1-alpha S) (Sporeophyte-specific EF-1-alpha)
 DE alpha)
 DE tEF-S.
 GN Porphyra purpurea.
 OS Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
 NCBI_TaxID:2787;
 QX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Avonport;
 RA MEDLINE:96309386; PubMed=8704161;
 RA Liu Q.Y., Baldauf S.L., Reith M.E.;
 RT "Elongation factor 1 alpha genes of the red alga Porphyra purpurea include a novel, developmentally specialized variant.";
 RT Plant Mol. Biol. 31:77-85(1996).
 CC -i- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF AMINOACYL-tRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN BIOSYNTHESIS.
 CC -i- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -i- DEVELOPMENTAL STAGE: EXPRESSED ONLY IN THE SPOROPHYTE, A SHELL-BORING, FILAMENTOUS PHASE.
 CC -i- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY. EF-TU/EF-1A SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
```

```

DR EMBL: U08841; AAA6179C.1; -.
DR HSSP: P07157; IALP.
DR InterPro: IPR004539; EFL_alpha.
DR InterPro: IPR000795; EE_GTPbind.
DR InterPro: IPR004160; EFTU_Cterm.
DR InterPro: IPR004161; EFTU_D2.
DR Pfam: PF00009; GTP_EFTU; 1.
DR Pfam: PF03143; GTP_EFTU_D2; 1.
DR Pfam: PF03144; GTP_EFTU_D3; 1.
DR TIGRFAMs: TIGR00483; EF-1_alpha; 1.
DR PROSITE: PS00301; EFACITOR_GTP; 1.
KW Elongation factor; protein biosynthesis; GTP-binding;
Multigene family.
FT NP_BIND 14 21 GTP (BY SIMILARITY).
FT NP_BIND 91 95 GTP (BY SIMILARITY).
FT NP_BIND 151 154 GTP (BY SIMILARITY).
FT SEQUENCE 515 AA; 56648 MW; EBAC3F4029F62350 CRC64;
Query Match 6.2%; Score 89.5; DB 1; length 515;
Best Local Similarity 24.7%; Pred. No. 2.1;
Matches 37; Conservative 20; Mismatches 56; Indels 37; Gaps 6;
QY 41 PPSLP-----LGGGSGSGEKVSVKMAAAW-----PSGSAPEAVTARLVGVWFSV 90
DB 261 PLRLPLQDVYKIGGTVGVGRVETGILKAGHQVTFEPAGKAAVEKSVEM-----HH 313
QY 91 ITGPGAVATSAGGESLKCEDLVGVGVICDKPK-----INDATOPPVNCTN-- 137
DB 314 TSVPQAIQDNGVFNKVLIVDKIRGD-VGGTKNDPPPTCEFLANVLIQDHKNRNGY 372
QY 138 -----YTAHVSCFPAPNITCKDSSGNETH 161
DB 373 TPVLDCHTAHIAKPFASILSKKDKRGKQTH 402
RESULT 4
ID LCT2_MOUSE STANDARD: PRT: 151 AA.
AC Q88803; O88804; O90WN3; Q92337;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Leukocyte cell-derived chemotaxin 2 precursor (Chondromodulin II)
DE (ChM-II).
GN LECT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANT VAL-123.
RC STRAIN-BAC/c; TISSUE=Liver;
RX MEDLINE=98382586; PubMed=9714793;
RA Yamagoe S., Watanabe I., Mizuno S., Suzuki K.;
RT "The mouse lect2 gene: cloning of cDNA and genomic DNA, structural
characterization and chromosomal localization.";
RL Gene 216:171-178(1998).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=Swiss Webster / NIH; TISSUE=Embryo, and Liver;
RX MEDLINE=99160594; PubMed=1005029;
RA Shukunami C., Kondo J., Wakai H., Takahashi K., Inoue H., Kamizono A.,
RA Hiraki Y.;
RT "Molecular cloning of mouse and bovine chondromodulin-II cDNAs and the
growth-promoting actions of bovine recombinant protein.";
RL J. Biochem. 125:436-442(1999).
CC -!- FUNCTION: Has a neutrophil chemotactic activity. Also a positive
regulator of chondrocyte proliferation.
-!- SUBCELLULAR LOCATION: Secreted.
-!- ALTERNATIVE PRODUCTS:
Event-Alternative splicing: Named isoforms=2;
Name=1; Synonyms=LECT2;
Name=2; Synonyms=LECT2Q;
IsoId=O88803.1; Sequence=Displayed;
IsoId=O88803.2; Sequence=VSP_003051;
-!- TISSUE SPECIFICITY: Highly expressed in liver and weakly in
testis. Not expressed in heart, brain, spleen, lung, skeletal
muscle and kidney.
-!- SIMILARITY: BELONGS TO THE LECT2 / MIM-1 FAMILY.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation. -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
```

```

DR EMBL: AB009687; BAA33383.1; -.
DR EMBL: AB009688; BAA33384.1; -.
DR EMBL: AB009689; BAA33385.1; -.
DR EMBL: AB009689; BAA33386.1; -.
DR EMBL: AF035161; AAF13302.1; -.
DR MGD: MGI:1278342; Lect2.
KW Chemotaxis; Signal; Alternative splicing.
FT SIGNAL 1 18 BY SIMILARITY.
FT CHAIN 19 151 LEUKOCYTE CELL-DERIVED CHEMOTAXIN 2.
FT VARSPLIC 98 151 FCVAFYDKPKYKSGIKKGLTLLPQIKYGIQSHVH
VENCDSDDPTAYL -> QRQAHTTINVTCTYWDKIQIPR
PIRELCONFLH (in isoform 2).
FT VARIANT 129 129 I -> V.
FT SEQUENCE 151 AA; 16405 MW; 1BAF444046B7AE8E CRC64;
Query Match 6.0%; Score 86.5; DB 1; Length 151;
Best Local Similarity 24.8%; Pred. No. 0.89;
Matches 29; Conservative 9; Mismatches 52; Indels 27; Gaps 4;
QY 78 TARLVGVWFSVVTGPGAVATISAGGESLKCEDLVGVGVICDKPKINDATQPPVNCTN 137
DB 4 TTILISAALLSSALAGPWANICASKSSNEIRTCDSYGGQYSAQ-----RIQR----- 51
QY 138 YTAHVSCFPAPNITCKDSSGNETHFTGNEVGFPPISCRNV-----NGYSYKV 185
DB 52 ---H---HPGVDVLCDSGVVYAPFTGKIVGQEKPYRNKNAINDGIRLSGRGFCVKI 102
RESULT 5
ADH2_KLULA
ID ADH2_KLULA STANDARD: PRT: 348 AA.
AC P49383;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Alcohol dehydrogenase II (EC 1.1.1.1).
GN ADH2.
OS Kluyveromyces lactis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=28985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS 2359 / IFO 1267 / NRRL Y-1140;
RX MEDLINE=9226769; PubMed=1588917;
RA Shain D.H., Salvadore C., Denis C.L.;
RT "Evolution of the alcohol dehydrogenase (ADH) genes in yeast:
characterization of a fourth ADH in Kluyveromyces lactis.";
RL Mol. Gen. Genet. 232:479-488(1992).
CC -!- CATALYTIC ACTIVITY: An alcohol + NAD(+) = an aldehyde or ketone +
NADH.
```

```

CC -|- COFACTOR: Binds 2 zinc ions per subunit (By similarity).
CC -|- SUBUNIT: Homotetramer (By similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -|- SIMILARITY: Belongs to the zinc-containing alcohol dehydrogenase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X64397; CA445739.1; -.
CC FIRM: S20911; S20911.
CC InterPro: IPR002328; ADH_zinc.
CC InterPro: IPR002085; ADH_zn_family.
CC Pfam: PF00107; ADH_zinc_N; 1.
CC PROSITE: PS00059; ADH_ZINC; 1.
CC Oxidoreductase; Zinc; Metal-binding; NAD: Multigene family.
CC METAL 44 44 ZINC 1 (CATALYTIC) (BY SIMILARITY).
CC METAL 67 67 ZINC 1 (CATALYTIC) (BY SIMILARITY).
CC METAL 98 98 ZINC 2 (BY SIMILARITY).
CC METAL 101 101 ZINC 2 (BY SIMILARITY).
CC METAL 104 104 ZINC 2 (BY SIMILARITY).
CC METAL 112 112 ZINC 2 (BY SIMILARITY).
CC METAL 154 154 ZINC 1 (CATALYTIC) (BY SIMILARITY).
CC SEQUENCE 348 AA; 37097 MW; F3B64AE1F520689C CRC64;
CC
CC Query Match: 5.9%; Score 85.5; DB 1; Length 348;
CC Best Local Similarity 19.9%; Pred. No. 2.9;
CC Matches 65; Conservative 35; Mismatches 111; Indels 115; Gaps 13;
CC
QY 19 NTRDGG---TGLPMRGPPKNAALPFSLPGGSGSGSERVSVSKMAAN----- 67
DB 37 NVKYGWCHTDLHAWGDWP-----LPTKLPLV-GGHEGAGVYVAMGENYKGMIIIGLFA 91
QY 68 -----PSGPSAPEAVT 75
DB 92 KWLNGSCMSCEYCELSNESNCPDADLSGYTHDGSFQYATADAVQAARIPKGTU-AEVAP 151
QY 79 ARLVGV-----LMFVSVTGPGWAVATSAGGESLKCEDLV----- 115
DB 152 ILCAGTVVYKALKSADLKAGDVAISGACGGGSGIAIOYAKAMGYRVGLTGTCAEKALF 211
QY 116 ----GQYICKPKINGATCEPVNCTNYTAH-----VSCFPAINITCKDSSGNETHTCN 155
DB 212 KELGGYFVDYAVSKDITREIVDATNGAHGV-NVSVSEFAI-----ECSTNYVRSNG 263
QY 166 EVGFFKPIPCRNVNGYSYKVAVALSLFLQWLG--AD-RFYVIGYFALGL-KFCTVGFQGLG 222
DB 266 VLVGLPRDAKCKSDVFTQVVKSVSVGVGNKRAOTREALDFANGLV-HAPIKIVGLS 324
QY 223 SLIDFI--LLSMGIVGPSOSSVIID 246
DB 325 ELADVYDKMYKGEIVG-----RYVVD 345
CC
RESULT 6
LAMP_RAT STANDARD; PRT; 338 AA.
AC Q62813;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Limbic system-associated membrane protein precursor (LSAMP).
GN LSAMP OR LAMP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID-10116;
RN [1]

```

```

PP SEQUENCE FROM N.A., AND SEQUENCE OF 29-49.
PC TISSUE=Hippocampus;
RX MEDLINE=95374785; PubMed=7646886;
RA Pimenta A.F.; Zhukareva V.; Barbe M.F.; Reinoso B.S.; Grimley C.;
RA Henzel W.; Fischer I.; Levitt P.;
RT "The limbic system-associated membrane protein is an Ig superfamily
RL member that mediates selective neuronal growth and axon targeting.";
CC -|- FUNCTION: MEDIATES SELECTIVE NEURONAL GROWTH AND AXON TARGETING.
CC CONTRIBUTES TO THE GUIDANCE OF DEVELOPING AXONS AND REMODELING OF
CC MATURE CIRCUITS IN THE LIMBIC SYSTEM. ESSENTIAL FOR NORMAL GROWTH
CC OF THE HIPPOCAMPAL MOSSY FIBER PROJECTION.
CC -|- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -|- TISSUE SPECIFICITY: EXPRESSED MOSTLY BY NEURONS COMPRISING LIMBIC-
CC ASSOCIATED CORTICAL AND SUBCORTICAL REGIONS THAT FUNCTION IN
CC COGNITION, EMOTION, MEMORY, AND LEARNING.
CC -|- DEVELOPMENTAL STAGE: FIRST DETECTED AT E15-16, AT STAGE E20 IT IS
CC DETECTED IN PRESUMPTIVE CORTEX, MEDIAL LIMBIC AREAS OF THE
CC THALAMUS AND HYPOTHALAMUS. IN THE ADULT, IT IS FOUND IN
CC HYPOTHALAMUS, PERIRHINAL CORTEX, AMYGDALA AND MEDIAL THALAMIC
CC REGION.
CC -|- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON
CC SUBFAMILY.
CC -|- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaborator
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: J31554; AAA86120.1; -.
CC InterPro: IPR007110; Ig_Like.
CC InterPro: IPR003598; Ig_C2.
CC InterPro: IPR003006; Ig_MHC.
CC Pfam: PF00047; Ig_3.
CC SMART: SM00408; IGC2; 2.
CC PROSITE: PS50835; IG_LIKE; 3.
CC Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
CC Repeat; Signal.
CC SIGNAL 1 28
CC CHAIN 29 315
CC
CC LIMBIC SYSTEM-ASSOCIATED MEMBRANE
CC PROTEIN.
CC REMOVED IN MATURE FORM (POTENTIAL).
CC IG-LIKE C2-TYPE 1.
CC IG-LIKE C2-TYPE 2.
CC IG-LIKE C2-TYPE 3.
CC POTENTIAL.
CC POTENTIAL.
CC POTENTIAL.
CC N-LINKED (GLCNAC) (POTENTIAL).
CC N-LINKED (GLCNAC) (POTENTIAL).
CC N-LINKED (GLCNAC) (POTENTIAL).
CC N-LINKED (GLCNAC) (POTENTIAL).
CC N-LINKED (GLCNAC) (POTENTIAL).
CC N-LINKED (GLCNAC) (POTENTIAL).
CC N-LINKED (GLCNAC) (POTENTIAL).
CC GPI-ANCHOR (POTENTIAL).
CC SEQUENCE 338 AA; 37324 MW; 0B76AFDD68A39BB6 CRC64;
CC
CC Query Match: 5.9%; Score 85; DB 1; Length 338;
CC Best Local Similarity 27.7%; Pred. No. 3.1;
CC Matches 36; Conservative 15; Mismatches 47; Indels 32; Gaps 7;
CC
QY 101 SAGGEESLKCEDLVKVG-----QYICKPKINDATCEPVNCTNYTAHVSCFPAINITCKDSS 156
DB 230 TTGROASLKCEASAVPADPFDWYRDDTRINSANGLEIKSTE---GQSLLIVTNTV-EEHY 285
QY 157 GNETHFTGNEVG-----FFKPIPCRNVNGYSYKVAVALSLFLQWLGADRFYGLYPALG 209
RN [1]

```

```

FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 163 183 POTENTIAL.
FT TRANSMEM 214 234 POTENTIAL.
FT TRANSMEM 256 276 POTENTIAL.
FT TRANSMEM 290 310 POTENTIAL.
FT TRANSMEM 341 361 POTENTIAL.
FT TRANSMEM 369 389 POTENTIAL.
FT TRANSMEM 414 434 POTENTIAL.
FT TRANSMEM 440 460 POTENTIAL.
SQ SEQUENCE 487 AA; 52194 MW; 3572502DB6ACD987 CRC64;

Query Match 5.8%; Score 83.5; DB 1; Length 487;
Best Local Similarity 26.5%; Pred. NO. 6.5;
Matches 27; Conservative 16; Mismatches 40; Indels 19; Gaps 5;

QY 159 ETHFTGNEGVCFKPISCRVNNGYSYKVAVALSLFLGNLGNADREYLGYPALGLLKFTVG 218
DB 8 DERLTREDTGKHYKGLHSROLMIALGATGTGFLG--AGRIASAGPGL---FLVYGI 61
QY 219 CGTGLSLDIFILISMQVVG-----PSDGS--SYIIDYGTGL 252
DB 62 CGI-----FVFLTRALGELVLRHPSSGSFVSAREFYGEKV 98

RESULT 2
ADH_KILUMA STANDARD: PNT; 348 AA.
AC Q07288;
CT 01-FEB-1995 (Rel. 31, Created)
CT 01-FEB-1995 (Rel. 31, Last sequence update)
CT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Alcohol dehydrogenase 1 (EC 1.1.1.1).
GN ADH1
CS Kluyveromyces marxianus (Yeast) (Kluyveromyces fragilis).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
NCBI_TaxID=4911;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12424;
RX MEDLINE=93250057; PubMed=8485163;
LA Ladiere J.M., Delcour J., Vandenhaute J.;
RT "Sequence of a gene coding for a cytoplasmic alcohol dehydrogenase
RT from Kluyveromyces marxianus ATCC 12424.";
RI Biochim. Biophys. Acta 1173:99-101(1993).
CC -!- CATALYTIC ACTIVITY: An alcohol + NAD(+) -> an aldehyde or ketone +
CC NADH.
CC -!- COFACTOR: Binds 2 zinc ions per subunit (By similarity).
CC -!- SUBUNIT: Homotetramer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the zinc-containing alcohol dehydrogenase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcements
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X60224; CAA42785.1; -.
CC PIR: S32521.
CC DR InterPro: IPR002328; ADH_zinc.
CC DR InterPro: IPR002085; ADH_zn_family.
CC DF Pfam: PF00107; ADH_zinc_N; 1.
CC DF PROSITE: PS00059; ADH_ZINC; 1.
CC KP Oxidoreductase; Zinc; Metal-binding; NAD: Multigene family.
CC FT METAL 44 44 ZINC 1 (CATALYTIC) (BY SIMILARITY).
CC FT METAL 67 67 ZINC 1 (CATALYTIC) (BY SIMILARITY).
CC FT METAL 98 98 ZINC 2 (BY SIMILARITY).
CC FT METAL 101 101 ZINC 2 (BY SIMILARITY).
CC FT METAL 104 104 ZINC 2 (BY SIMILARITY).

```

```

FT METAL 112 112 ZINC 2 (BY SIMILARITY).
FT METAL 154 154 ZINC 1 (CATALYTIC) (BY SIMILARITY).
SQ SEQUENCE 348 AA: 37158 MW: A75D2EBE82E35BD CRC64;

Query Match 5.8%; Score 83; Dh 1; Length 348;
Best Local Similarity 20.8%; Pred. No. 4.9;
Matches 65; Conservative 38; Mismatches 107; Indels 102; Gaps 14;

QY 19 NTRDQ---TGYPVMPGPKNIALLPSPULPLGGSGSGGEKVSYSKMAAAPSG-----70
Db 37 NVKYSVCHTDLHAWGDWP-----LDTKPLV-GEHAGAGIVVANGENVTEIGDYAGI 9;

QY 71 -----PSAPEA-----VTARLV-----82
Db 92 KWLNGSCMCECELSNCPKADLSGYTHDGSFOYATADAVOAAIPKKNVDLAEPV 151
QY 83 ----GV-----LWFVSVTGPGCATSVTSAGGEESLKCEDLV-----115
Db 152 ILCAGVTYVYKALSAHIKAGDWVATSGAGCGGLSLAIQYAKAMGYRVLGIDAGDEKAKLF 211
QY 116 ----GOYICKDKINDATPEVNCINVTYAHVSCFPAPNITCKDSGNETFTTGN---VGF 169
Db 212 KELGGEYFDFTKDKDWAIEATNGVAHAINVSSEAAISTSVLYTRNSGTVVLYGL 27;
QY 170 FPIPCRNNGVYKYVAVALSLFLGLWG--AD-RFYLVPAIQLLK--FCTVGFQGTGSL 224
Db 272 PRDAOCK--SDFNVQVKSISVGSVYVGNRADTREALDFSRGLVKAP-KILGLSELASV 329
QY 225 IDPILISMIOIVG 236
Db 330 YD-KMKVQGIQV 340

RESULT 9
CFAB_HUMAN STANDARD; PRI: 764 AA
AC P00751; O15006; Q29944; Q96HX6; Q9BTF5; Q9BX92;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement factor B precursor (EC 3.4.21.47) (C3/C5 convertase)
DE (Properdin factor B) (Glycine-rich beta glycoprotein) (GBG) (PB2).
GN BF.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ALLELES S; FA AND FB).
RX MEDLINE=91065702; PubMed=2249879;
RA Davrinche C., Abbai M., Clerc A.;
R "Molecular characterization of human complement factor B subtypes.";
RL Immunogenetics 32:309-312(1990).
RN [2]
RP SEQUENCE FROM N.A. (ALLELE S).
RC TISSUE=Liver;
RX MEDLINE=94237735; PubMed=8181962;
RA Mejia J.E., Jain I., de la Salle H., Hauptmann G.;
R "Human factor B. Complete cDNA sequence of the RF*S allele.";
RL Hum. Immunol. 39:49-53(1994).
RN [3]
RP SEQUENCE FROM N.A. (ALLELE S).
RC TISSUE=Liver;
RX MEDLINE=94041399; PubMed=8225386;
RA Schwaible W., Luettig B., Sokolowski T., Estalier C., Weiss E.H.,
R Meyer Zum Bueschenfelde K.-H., Whaley K., Dippold W.;
R "Human complement factor B: functional properties of a recombinant
R zymogen of the alternative activation pathway convertase.";
RL Immunobiology 188:221-232(1993).
RN [4]
RP SEQUENCE FROM N.A. (ALLELE S).
RX MEDLINE=94067177; PubMed=8247029;
RA Horiuchi T., Kim S., Matsumoto M., Watanabe I., Fujita S.,

```

```

RA Volanakis J.E.;
RT "Human complement factor B: cDNA cloning, nucleotide sequencing,
RT phenotypic conversion by site-directed mutagenesis and expression.";
RL Mol. Immunol. 30:1587-1592(1993).
RN [5]
RP SEQUENCE FROM N.A.
RA Rwen L., Dankers C., Baskin D., Faust J., Loretz C., Alcant M.E.,
RA Banta A., Swartzell S., Smith T.M., Spies T., Hood L.;
R "Sequence determination of 300 kilobases of the human class III MHC
RT locus.";
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins B.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Aleschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Djatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rata S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulláhy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.T., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
R "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [7]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RA Jaatinen T., Kanerva J., Poutanen K.E., Saarinen-Pihkala U.,
RA Lokki M.-L.;
R "Expression and alternative splicing of human factor B gene in
RT leukemic mononuclear cells.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE OF 26-764, PARTIAL SEQUENCE FROM N.A., AND CARBOHYDRATES.
RX MEDLINE=84161997; PubMed=6546754;
RA Mole J.E., Anderson J.K., Davidson E.A., Woods D.E.;
R "Complete primary structure for the zymogen of human complement
RT factor B.";
RL J. Biol. Chem. 259:3407-3412(1984).
RN [9]
RP SEQUENCE OF 260-764.
RX MEDLINE=83204002; PubMed=6342610;
RA Christle D.L., Gagnon J.;
R "Amino acid sequence of the Bb fragment from complement factor B.
RT Sequence of the major cyanogen bromide-cleavage peptide (CB-II) and
RT completion of the sequence of the Bb fragment.";
RL Biochem. J. 209:61-70(1983).
RN [10]
RP SEQUENCE OF 339-764 FROM N.A.
RX MEDLINE=83273641; PubMed=6308626;
RA Campbell R.D., Porter R.R.;
R "Molecular cloning and characterization of the gene coding for human
RT complement protein factor B.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:4464-4468(1983).
RN [11]
RP SEQUENCE OF 467-595 AND 752-764 FROM N.A.
RX MEDLINE=83039428; PubMed=6957884;
RA Woods D.E., Markham A.F., Ricker A.T., Goldberger G., Colten H.R.;
R "Isolation of cDNA clones for the human complement protein factor B,
RT a class III major histocompatibility complex gene product.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:5661-5665(1982).
RN [12]
RP SEQUENCE OF 16-259 FROM N.A.

```



```

RX MEDLINE-84158524; PubMed-6323161;
RA Morley B.J., Campbell R.D.;
RT "Internal homologies of the Ba fragment from human complement
RI component Factor B, a class III MHC antigen.";
RN EMBO J. 3:153-157(1984).
RN [13]
RP SEQUENCE OF 1-99 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE-87102880; PubMed-3643061;
RA Wu L.-C., Morley B.J., Campbell R.D.;
RT "Cell-specific expression of the human complement protein factor B
RI gene: evidence for the role of two distinct 5'-flanking elements.";
RN Cell 48:331-342(1987).
RN [14]
RP GLYCATION OF LYS-291.
RX MEDLINE-91174758; PubMed-2006911;
RA Nlemann M.A., Shown A.S., Miller E.J.;
RT "The principal site of glycation of human complement factor H.";
RN Biochem. J. 274:473-480(1991).
CC -!- FUNCTION: FACTOR B WHICH IS PART OF THE ALTERNATE PATHWAY OF THE
CC COMPLEMENT SYSTEM IS CLEAVED BY FACTOR D INTO 2 FRAGMENTS: BA AND
CC BB. BA, A SERINE PROTEASE, THEN COMBINES WITH COMPLEMENT FACTOR IS
CC TO GENERATE THE C3 OR C5 CONVERTASE. IT HAS ALSO BEEN IMPLICATED
CC IN PROLIFERATION AND DIFFERENTIATION OF PREACTIVATED B
CC LYMPHOCYTES, RAPID SPREADING OF PERIPHERAL BLOOD MONOCYTES.
CC STIMULATION OF LYMPHOCYTE BLASTOGENESIS AND LYSIS OF ERYTHROCYTES.
CC BA INHIBITS THE PROLIFERATION OF PREACTIVATED B LYMPHOCYTES.
CC -!- CATALYTIC ACTIVITY: Cleaves C3 in the alpha-chain to yield C3a and
CC C3b. Cleaves C5 in the alpha-chain to yield C5a and C5b. Both
CC cleavages take place at the C-terminal of an arginine residue.
CC -!- SUBUNIT: Monomer.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing: Named isoforms=2;
CC Name=1;
CC IsoId=P00751-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P00751-2; Sequence=VSP_005380, VSP_005381;
CC -!- POLYMORPHISM: TWO MAJOR VARIANTS, F AND S, AND 2 MINOR VARIANTS,
CC AS WELL AS AT LEAST 14 VERY RARE VARIANTS, HAVE BEEN IDENTIFIED.
CC -!- MISCELLANEOUS: FACTOR B IS A MAJOR HISTOCOMPATIBILITY COMPLEX
CC CLASS-III PROTEIN.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 3 Sushi (SCR) domains.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X72875; CAAS1389.1; -
DR EMBL: S67310; AAD13989.1; -
DR EMBL: L15702; AAA16820.1; -
DR EMBL: X00284; CAA25077.1; -
DR EMBL: AF019413; AAB67977.1; -
DR EMBL: BC004143; AAH04143.1; -
DR EMBL: BC007990; AAH07990.1; -
DR EMBL: AF349679; AAH30167.1; -
DR EMBL: K01566; AAA36225.2; -
DR EMBL: J00125; -; NOT_ANNOTATED_CDS.
DR EMBL: J00126; AAA36226.1; -
DR EMBL: J00185; AAA36219.1; ALT_SEQ.
DR EMBL: J00186; AAA36220.1; -
DR EMBL: M15082; AAA59625.1; -
DR PIR: S34075; BBHU.
DR HSSP: P20231; IAAO.
DR HEMO2S; S01196; -
DR SWISS-2DPAGE: P00751; HUMAN.
DR Slena-2DPAGE: P00751; -
DR Genew: HGNC:1037; BF.

```

```

DR MIM: 138470; -
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser. protease.Try.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR InterPro: IPR002035; VWFA.
DR Pfam: PF00084; sushi; 3.
DR Pfam: PF00089; trypsin; 1.
DR Pfam: PF00092; vwa; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00453; VWFADOMAIN.
DR SMART: SM00032; CCP; 3.
DR SMART: SM00020; Tryp_SPC; 1.
DR SMART: SM00327; VWA; 1.
DR PROSITE: PS50240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.

Query Match          5.8%; Score 83; DB 1; Length 764;
Best Local Similarity 24.1%; Pred. No. 12;
Matches 49; Conservative 21; Mismatches 71; Indels 62; Gaps 12;

QY 24 GTGLYPMRCPPKNNLALLPSPLLGK-----GSGSGEKVS----- 60
DE 2 GSNLSP-----QLCLMPFILLGSGVTTTPWSLARFQSGSLGVKIGSFRLLQEG 55
QY 61 SKMAAAPSG--PSAPEAVTARLVGLVFWVSVTGFWGAVATS-----AGGEESLKC-- 110
DB 56 QALEYVCSGFGYPVQVTRICR-----STGSNLIKTDQOKTYRKACRAIHGPR 105
QY 111 -EDLVGVYICKDKINDATQEPVNC-TNYTAHVSCTFPAPN-TCKDSS--GNFTHTGNE 166
DB 106 PHDFENGVEWPRSPYVNVSDSIFHCYDGYTLRGS-----NRTQVNGRWSGQTACDNG 161
QY 167 VGFFK----PISCRNVNGYSYKV 185
DB 162 AGYCSNPGIPIGTRKV-GSQYRL 183

RESULT 10
LAMP_HUMAN
ID LAMP_HUMAN STANDARD; PRT; 338 AA.
AC Q33449;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Limbic system-associated membrane protein precursor (LSAMP).
GN LSAMP OR LAMP.
OS Homo sapiens (Human).
OC Fukuyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-96235133; PubMed-8666243;
RA Pimente A.F., Fischer I., Levitt P.;
RT "cDNA cloning and structural analysis of the human limbic-system-
RI associated membrane protein (LAMP).";
RN Gene 170:189-195(1996).
CC -!- FUNCTION: MEDIATES SELECTIVE NEURONAL GROWTH AND AXON TARGETING.
CC CONTRIBUTES TO THE GUIDANCE OF DEVELOPING AXONS AND REMODELING OF
CC MATURE CIRCUITS IN THE LIMBIC SYSTEM. ESSENTIAL FOR NORMAL GROWTH
CC OF THE HYPOCAMPAL MOSSY FIBER PROJECTION (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- TISSUE SPECIFICITY: EXPRESSED ON LIMBIC NEURONS AND FIBER TRACTS
CC AS WELL AS IN SINGLE LAYERS OF THE SUPERIOR COLLICULUS, SPINAL
CC CHORD AND CEREBELLUM.
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON
CC SUBFAMILY.
CC -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U41901; AAC50569.1; .
DR PIR: J04776; J04776.
DR Genbank: HGRC:6705; LSAMP.
DR MIM: 603241; .
DR GO: GO:007399; P:neurogenesis; TAS.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003598; Ig-C2.
DR InterPro: IPR003006; Ig-MHC.
DR Pfam: PF00047; Ig: 3.
DR SMART: SM00408; IGC2; 2.
DR PROSITE: PS00835; IG-LIKE; 3.
KW Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
KW Repeat; Signal.
FT SIGNAL 1 28
FT CHAIN 29 315
FT LIMBIC SYSTEM-ASSOCIATED MEMBRANE
FT PROPEP 316 338
FT DOMAIN 29 122
FT DOMAIN 132 214
FT DOMAIN 219 304
FT DISULFID 53 111
FT DISULFID 153 197
FT DISULFID 239 290
FT CARBOHYD 40 40
FT CARBOHYD 66 66
FT CARBOHYD 136 136
FT CARBOHYD 148 148
FT CARBOHYD 279 279
FT CARBOHYD 287 287
FT CARBOHYD 300 300
FT CARBOHYD 315 315
FT LIPID 315 315
SQ SEQUENCE 338 AA: 37308 MW: 034552860FSD92F CRC64;
Query Match 5.7%; Score 82.5; DB 1; Length 338;
Best Local Similarity 29.6%; Pred. No. 5.2;
Matches 37; Conservative 14; Mismatches 47; Indels 27; Gaps 7;
QY 101 SAGGESLKCDLVKGV-----QVCKDPKINATQPVNCTNTHAVSCFPAPN:TCDDSS 156
DB 230 TTGQASLKCEASVAPDPDEFYRDTRDTSANGLKIKSTE---GQSSLTYNVT-EEHY 235
QY 157 GNETHFTGNEVG-----FFKPIPCRNWNGYSKVAVALSLFLGLGADRFY-GYPALG 209
DB 286 GNYTCVANKLGVTNATSLVIFPGSVRGING-SISLAVPL-----WL-----LAASLAC 333
QY 210 LKFC 214
DB 334 LLSKC 338
RESULT 11
VG32-BPM02
ID VG32-BPM02 STANDARD; PRT: 223 AA.
AC 064226;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gene 32 protein (GP32).
GN 32.
OS Mycobacteriophage D29.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=28369;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98300335; PubMed=9636706;
RA Ford M.E., Sarkis G.J., Belanger A.E., Hendrix R.W., Hatfull G.F.;
RT "Genome structure of mycobacteriophage D29: implications for phage

RT evolution.";
RL J. Mol. Biol. 279:143-164(1998).
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF022214; AAC18473.1; .
DR PIR: F72803; F72803.
SQ SEQUENCE 223 AA: 21822 MW: 33CD0DC310038AD4 CRC64;
Query Match 5.6%; Score 80.5; DB 1; Length 223;
Best Local Similarity 30.7%; Pred. No. 4.8;
Matches 27; Conservative 8; Mismatches 36; Indels 17; Gaps 1;
QY 29 PMRGPFKNLALLPFLP-----LLGGSGSGGKVSVMKMAAWPSPGSAPEAV-A 79
DB 37 PVLTPVTAVGAYTYNIPAAQAEFIDVLLGAGGGGQ-----MGSATANGGGGFGGSWV-A 91
QY 80 RL---VGVLFVSVYTGPCWCAVATSAGG 104
DB 92 TLRRGVDIPWAVTQITGVIGAGGTAGPG 119
RESULT 12
ENT1_HUMAN
ID ENT1_HUMAN STANDARD; PRT: 455 AA.
AC Q99808;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Equilibrative nucleoside transporter 1 (Equilibrative
DE nicotinylmercapturine riboside-sensitive nucleoside transporter)
DE (Equilibrative NBMPR-sensitive nucleoside transporter) (Nucleoside
DE transporter, es-type).
GN SLC29A1 OR ENT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-21.
RX TISSUP-Placenta;
MEDLINE=97140266; PubMed=8986748;
RA Griffiths M., Beaumont N., Yao S.Y.M., Sundaram M., Bouman C.E.,
RA Davies A., Kwong F.Y.P., Coe I., Cass C.E., Young J.D., Baldwin S.A.;
RA "Cloning of a human nucleoside transporter implicated in the cellular
RA uptake of adenosine and chemotherapeutic drugs.";
RA Nat. Med. 3:89-93(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUP-Jejunum, and Small intestine;
RA Lum P.Y., Ngo L.Y., Bakken A.H., Unadkat J.D.;
RA "Critical structural determinants for high affinity binding of
RA nucleosides to the equilibrative NBMPR-sensitive nucleoside
RA transporter (es) cloned from the human jejunum.";
RA Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases
CC -1- FUNCTION: MEDIATES BOTH INFILUX AND EFFLUX OF NUCLEOSIDES ACROSS
CC THE MEMBRANE (EQUILIBRATIVE TRANSPORTER). IT IS SENSITIVE (ES) TO
CC LOW CONCENTRATIONS OF THE INHIBITOR NITROBENZYLMECAPTOPURINE
CC RIBOSIDE (NBMPR) AND IS SODIUM-INDEPENDENT. IT HAS A HIGHER
CC AFFINITY FOR ADENOSINE, INHIBITED BY DIPYRIDAMOLE AND DILAZEP
CC (ANTICANCER CHEMOTHERAPEUTICS DRUGS).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN HEART, BRAIN, MAMMARY GLAND,
CC ERYTHROCYTES AND PLACENTA, AND ALSO IN FETAL LIVER AND SPLEEN.
CC -1- PTM: GLYCOSYLATED.
CC -1- SIMILARITY: BELONGS TO THE SLC29A FAMILY OF TRANSPORTERS.
CC -----


```

CC -!- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a
CC fatty acid anion.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE ARHYDROLASE SUPERFAMILY. LIPASE FAMILY.
CC -!- SIMILARITY: Contains 1 Pfam domain.
CC -!- DATABASE: NAME-Worthington enzyme manual;
CC WWW="http://www.worthington-biochem.com/manual/L/P/L.html".
CC PDB: 1FTH: 07-DEC-96.
CC InterPro: IPRO00734; Lipase.
CC InterPro: IPRO01024; Lipoxigenase_LH2.
CC InterPro: IPRO00379; Ser_cstrs_site.
CC Pfam: PF00151; Lipase; 1.
CC Pfam: PF01477; PLAT; 1.
CC PRINTS: PR00821; TAGLIPASE.
CC SMART: SM00308; LH2; 1.
CC PROSITE: PS00120; LIPASE_SER; 1.
CC PROSITE: PS0095; PLAT; 1.
CC Hydrolyase; Lipid degradation; Pancreas; Glycoprotein; 3D-structure.
KW DOMAIN 339 450
FT ACT_SITE 153 153 CHARGE RELAY SYSTEM.
FT ACT_SITE 177 177 CHARGE RELAY SYSTEM.
FT ACT_SITE 264 264 CHARGE RELAY SYSTEM.
FT DISULFID 4 10
FT DISULFID 91 102 IN ISOMER 1.
FT DISULFID 91 104 IN ISOMER 2.
FT DISULFID 238 262
FT DISULFID 286 297
FT DISULFID 300 305
FT DISULFID 434 450
FT CARBOHYD 157 167 N-LINKED (GLCNAC...)
SQ SEQUENCE 450 AA; 50084 MW; 76E13BB3B451PCE CRC64;

Query Match 5.6%; Score 80; DB 1; Length 450;
Best Local Similarity 20.2%; Pred. No. 12;
Matches 47; Conservative 30; Mismatches 96; Indels 60; Gaps 10;

QY 1 MHILKSPNVIP-----RAHGQKNTRRDQ-----TCLYPMRGPFKNALL---40
DB 133 VEVLKSSLGSPSNVHVIGHSLGSHAAGEARTNGTIERITCLDPAEPCFGTPE:VRL 192
QY 41 -----PFSILPLGGSGGSEKVSXMAAWPSG-----PSAPEAVTARLVGV 34
DB 193 DPSDAKFDVVIHTDAAPIIPNLGFGMSQTVGHLPF---FPNGGKOMPCCOKNLSOIVEI 244
QY 85 LWFVSVTTGPGAVATSAGSESLCEDLKVCQYICKDKPKINDATQEPVNCINYAHVS- 143
DB 250 -----DGIW-----EGTRQFVACNHLSRYKYA-CSLAPDGGPAGPDCSYNVFTAN 295
QY 144 -GFPAPNITCKDSSGNETHFTG-----NEVGFFKP-SCRNVNGYSYKVAVALS 190
DB 296 KCFPCPSECGPQMGHYADRFPGKTNVGSQVYINTGFSNFAWRKYKVSVTLS 348

RESULT 15
OAL_LYMST
ID OAL_LYMST STANDARD; PRT: 638 AA.
AC 077408;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Octopamine receptor 1 (OAL).
OS Lymnaea stagnalis (Great pond snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
OC Lymnaeidae; Lymnaeidae; Lymnaea.
OX NCBI_TaxID=6523;
RN [1]
RP SEQUENCE FROM N.A.. AND CHARACTERIZATION.
RC TISSUE= CNS;
RX MEDLINE=97347296; PubMed=9203635;
RA Gerhardt C.C., Baker R.A., Plek R.J., Vreugdenhil E.,
RA Leyssen J.E., van Heerikhuizen H.;
RT *Molecular cloning and pharmacological characterization of a molluscan
RT octopamine receptor*;
```

```

RL Mol. Pharmacol. 51:293-300(1997).
CC -!- FUNCTION: G-PROTEIN COUPLED RECEPTOR FOR OCTOPAMINE (OA), WHICH IS
CC A NEUROTRANSMITTER, NEUROHORMONE, AND NEUROMODULATOR IN
CC INVERTEBRATES. ACTIVATION OF THIS RECEPTOR BY OCTOPAMINE INDUCES
CC AN INCREASE IN BOTH INOSITOL PHOSPHATES AND CYCLIC AMP. THE
CC COUPLING TO ADENYLYL CYCLASE SEEMS TO BE LESS EFFICIENT THAN THE
CC COUPLING TO PHOSPHOLIPASE C. THE RANK ORDER OF POTENCY FOR
CC AGONISTS IS P-SYNEPHRINE >= CLONIDINE > P-OCTOPAMINE =
CC XYLOMETHAZOLINE = PHENYLEPHRINE = OXYMETAZOLINE > B-HT920 >
CC SEROTONIN = P-TYRAMINE > EPINEPHRINE > NOREPINEPHRINE >
CC METHOXAMINE = DOPAMINE - HISTAMINE. FOR ANTAGONISTS, THE RANK
CC ORDER IS YOHIMBINE > CHLOPHAZAMINE / SPIPERONE > PHENTOLAMINE >
CC MANSERINE > RAUWOLESCEINE > PRAZOSIN > ALPRENOLOL / PROPRANOLOL >
CC PINDOLOL.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE CENTRAL NERVOUS SYSTEM.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U62771; AAC61296.1; -.
CC HSSP: 229274; 1MM5;
CC InterPro: IPR000276; GPCR_Rhodpsn.
CC Pfam: PF00001; 7tm1.1;
CC PRINTS: PR00237; GPCR_RHODOPSIN.
CC PROSITE: PS00237; G-PROTEIN RECF_F1_1; 1.
CC PROSITE: PS0262; G-PROTEIN RECF_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein.
KW DOMAIN 1 28
FT TRANSMEM 29 53 POTENTIAL.
FT DOMAIN 54 64 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 65 87 POTENTIAL.
FT DOMAIN 88 102 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 103 124 POTENTIAL.
FT DOMAIN 125 147 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 148 167 POTENTIAL.
FT DOMAIN 168 239 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 240 259 POTENTIAL.
FT DOMAIN 260 520 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 521 545 POTENTIAL.
FT DOMAIN 546 551 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 552 575 POTENTIAL.
FT DOMAIN 576 638 CYTOPLASMIC (POTENTIAL).
FT DISULFID 101 230 BY SIMILARITY.
FT CARBOHYD 178 178 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 207 207 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 215 215 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 638 AA; 70000 MW; 65FA928B5C01D34F CRC64;

Query Match 5.6%; Score 80; DB 1; Length 638;
Best Local Similarity 24.8%; Pred. No. 18;
Matches 31; Conservative 18; Mismatches 56; Indels 20; Gaps 4;

QY 149 NITCKD-----SSGNETHFTGNEVGFFKPISCRNVNGYSYKVAVALSFLGLWGA 198
DB 69 NLACADLLGLVLVLPFSAVNEI-----KDWIFGHVWCQVWLAVDVMICTASILNCCISL 124
QY 199 DRFY-----LGPALGLLKFCVTGFCGICSLIDFLISMQIVGSPDGSYSYIDYGTGLT 253
DB 125 DYLAITPIRYPGLMSAKRAKTLVAGVM-LFSFVICPPLIGNMDGGIMDYGTTAT 183
QY 254 RLSIT 258
DB 184 PIPVT 188

Search completed: September 26, 2003, 17:25:45
```


GenCore version 5.1.6
Copyright: (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2003, 17:28:58 ; Search time 16.7106 Seconds
(without alignments)
313.452 Million cell updates/sec

Title: US-09-852-100A-2_COPY_185_217

Perfect score: 33

Sequence: 1 VAVALSFLNGLGADRYLGYPA:GLIKFVVG 33

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1107463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- A_Geneseq_19Jun03.*
- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
 - 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
 - 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
 - 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
 - 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
 - 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
 - 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
 - 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
 - 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
 - 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
 - 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
 - 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
 - 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
 - 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
 - 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
 - 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
 - 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
 - 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
 - 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
 - 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
 - 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
 - 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
 - 23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
 - 24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	33	100.0	269	AAW94291	Human beta-amyloid
2	33	100.0	269	AA197059	Human beta-amyloid
3	33	100.0	269	AAE33877	Human BPP-1 protein
4	17	51.5	139	AA123358	Human 5' ES: socie
5	9	27.3	225	ABR11574	Human secreted pro
6	8	24.2	24	ABB77688	New peptide vector
7	8	24.2	27	ABH1176	Signal sequence ba
8	8	24.2	27	ABG78990	Cell penetrating p
9	8	24.2	27	AAE23685	Fluorescently labe

10	8	24.2	27	23	ABB77687	New peptide vector
11	8	24.2	27	23	AAU78348	Signal sequence ba
12	8	24.2	27	24	ABG75506	Signal sequence ba
13	8	24.2	27	24	ABB82542	Signal sequence ba
14	8	24.2	178	22	ABG65236	Drosophila melanog
15	8	24.2	221	21	AAV70761	Human beta-amyloid
16	8	24.2	221	21	ABR47818	Human secreted pro
17	8	24.2	221	24	ABR48493	Human Amyloid App
18	8	24.2	221	24	ABR00112	Human gene 102 enc
19	8	24.2	222	20	AAV27658	Human secreted pro
20	8	24.2	230	21	ABO8645	DNA encoding a mur
21	8	24.2	247	21	ABO8641	A human TANGO 223
22	8	24.2	247	21	ABO8656	A human TANGO 223
23	8	24.2	247	21	ABO8657	A human TANGO 223
24	8	24.2	247	21	ABO8658	A human TANGO 223
25	8	24.2	284	22	ABB59014	Drosophila melanog
26	8	24.2	358	23	ABP62954	Human polypeptide
27	7	21.2	122	23	ABP63670	Human ORF40. Homc
28	7	21.2	173	21	AAV9435	Human PRO1489 (UNC
29	7	21.2	173	22	AAE66184	Protein of the inv
30	7	21.2	206	22	ABG22057	Novel human diagnc
31	7	21.2	223	13	AAE20227	Human claudin-D17
32	7	21.2	224	23	AAE22228	Human secreted pro
33	7	21.2	367	21	ABP39030	Human secreted pro
34	7	21.2	466	23	ABP26809	Streptococcus pol
35	7	21.2	489	21	ABR42550	Human ORF2314
36	7	21.2	489	21	AAI44453	Human carbamoyl pr
37	7	21.2	489	22	ABG92615	Human protein sequ
38	7	21.2	459	24	AAE33210	Human mitochondria
39	7	21.2	755	24	ABJ25413	Aspergillus fumiga
40	7	21.2	864	24	ABJ26013	Aspergillus fumiga
41	7	21.2	869	23	ABJ10922	Human secreted pro
42	7	21.2	1240	22	AAU56553	Propionibacterium
43	6	18.2	9	20	AAI10404	T cell epitope/MHC
44	6	18.2	9	23	ABG80086	MHC class I molecu
45	6	18.2	11	21	AAO80587	Peptide identific

ALIGNMENTS

RESULT 1					
AAW94291					
ID	AAW94291	standard; Protein; 269 AA.			
XX	AAW94291;				
AC	AAW94291;				
XX	27-APR-1999	(first entry)			
DT	Human beta-amyloid peptide-binding protein (BBP).				
DE	Human beta-amyloid peptide-binding protein; BBP: beta-amyloid protein; BAP;				
XX	human; Alzheimer's disease.				
XX	Homo sapiens.				
OS					
XX	Key	Location/Qualifiers			
EH	Region	68..269			
FT		/note= "specifically claimed fragment having			
FT		beta-amyloid peptide binding activity"			
XX					
XX	WO9846636-A2.				
PN					
XX	22-OCT-1998.				
XX					
XX	14-APR-1998;	98WO-US07462.			
XX					
P3	16-APR-1997;	97US-0064583.			
XX					
XX	(AMHP) AMERICAN HOME PROD CORP.				
PA	Bard JA, Jacobsen JS, Kajkowski EM, Ozenberger BA;				
XX	Walker SG;				
P1					

```
XX
DR WPI: 1999-080736/27.
DR N-PSDB; AAX05735.
XX
PT Polynucleotide encoding beta-amyloid peptide binding protein - used
PT to identify inhibitors of beta-amyloid peptide for treating
PT Alzheimer's disease
XX
PS Claim 7; Pages 43-44; 59pp; English.
XX
CC The present sequence represents a beta-amyloid peptide binding protein
CC (BBP). The polynucleotide comprising the entire BBP nucleotide sequence
CC of clone BBP1-f1 is deposited under the accession number: AACC 94517. The
CC polynucleotide comprising a fragment of BBP (nucleotides 202-807 of the
CC full length BBP) of clone PK196 is deposited as ATCC 98399. Host cells
CC transformed with a vector comprising the BBP nucleic acid are used for
CC the recombinant production of the protein. The protein can be used in a
CC method for diagnosing a disease characterised by aberrant expression of
CC human beta-amyloid protein (BAP). The protein can also be used in a
CC method for screening for compounds which regulate expression of a BAP
CC binding protein. The proteins, antibodies and identified compounds can be
CC used in the treatment or prevention of Alzheimer's disease.
XX
SQ Sequence 269 AA:
XX
Query Match 100.0%; Score 33; DB 20; Length 269;
Best Local Similarity 100.0%; Pred. No. 3.9e-24;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VAVALSLFLGWLGLGADRFYLGYPALGLLKFCVTG 33
Db 185 VAVALSLFLGWLGLGADRFYLGYPALGLLKFCVTG 217
RESULT 2
AAY70759
ID AAY70759 standard; Protein: 269 AA.
XX
AC AAY70759;
XX
DT 24-JUL-2000 (first entry)
XX
DE Human beta-amyloid peptide (BAP) binding protein, BBP1.
XX
KW Beta-amyloid peptide binding protein; BBP; BAP; tumour; suppressor;
KW G-protein coupled receptor; GPCR; integral membrane protein; antigen;
KW neuronal cell; nonhuman primate; NHP; G-protein signalling pathway;
KW apoptosis; immunogen; therapeutic; treatment; prevention; diagnostic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 177..198
FT /label= Transmembrane_domain_1
FT Domain 199..201
FT /label= DRF_motif
FT /note= "Substitution of the Arg abrogates protection"
FT Domain 213..238
FT /label= Transmembrane_domain_2
XX
PN WO2000022125-A2.
XX
PD 20-APR-2000.
XX
PF 13-OCT-1999; 99WO-US21621.
XX
PR 13-OCT-1998; 98US-0104104.
XX
PA (AMHP ) AMERICAN HOME PROD CORP.
XX
XX Ozenberger BA, Kajkowski EM, Lo CF;
XX WPI: 2000-317982/27.
```

```
DR N-PSDB; AAZ52369.
XX
PT Novel G-protein-coupled receptor-like proteins and polynucleotides
PT useful for regulating apoptosis, comprises integral membrane protein
PT traversing the membrane twice.
XX
PS Example 1; Page 62-63; 68pp; English.
XX
CC The present sequence is the beta-amyloid peptide (BAP) binding protein-1
CC (BBP1). It is an integral membrane protein, that traverse the membrane
CC twice. It is related to G protein-coupled receptor (GPCR) protein
CC superfamily. It interacts with G-alpha proteins and regulates the
CC activity of G-protein signalling pathways. BBP genes are widely expressed
CC in neuronal cells of nonhuman primate (NHP) brain and overexpressed in
CC some tumours. It functions as a suppressor of apoptosis induction. BBP
CC proteins are used as immunogens to raise antibodies, useful as
CC therapeutics and as antigens in solid phase assays. They are also useful
CC as reagents to identify molecules which effect the interaction of BBP and
CC diseases associated with apoptosis. The polynucleotides are useful for
CC diagnostics.
CC Note: In claim 5, the patent claims an amino acid sequence from figure 2.
CC However, figure 2 does not contain any sequence. It is inferred from the
CC disclosure that the figure 2 sequence refers to BBP1 protein, shown in
CC this sequence.
XX
SQ Sequence 269 AA:
XX
Query Match 100.0%; Score 33; DB 21; Length 269;
Best Local Similarity 100.0%; Pred. No. 3.9e-24;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VAVALSLFLGWLGLGADRFYLGYPALGLLKFCVTG 33
Db 185 VAVALSLFLGWLGLGADRFYLGYPALGLLKFCVTG 217
RESULT 3
AAE33877
ID AAE33877 standard; Protein: 269 AA.
XX
AC AAE33877;
XX
DT 02-MAY-2003 (first entry)
XX
DE Human; BBP-1 protein.
XX
KW Human; beta-amyloid peptide-binding protein; BAP; Abeta; betaAP; BBP;
KW Alzheimer's disease; AD; transgenic; transgenic animal; gene therapy;
KW neuroprotective; nootropic.
XX
OS Homo sapiens.
XX
FN WO200290499-A2.
XX
PD 14-NOV-2002.
XX
PF 06-MAY-2002; 2002WO-US14223.
XX
PR 09-MAY-2001; 2001US-0852100.
XX
PA (AMHP ) WYETH.
XX
PI Ozenberger BA, Bard JA, Kajkowski EM, Jacobsen JS, Walker SG;
PI Sofia HJ, Howland DS;
XX
XX WPI: 2003-120537/11.
DR N-PSDB; AAD51940.
XX
XX New human beta-amyloid peptide-binding protein, useful for diagnosing
XX and/or treating diseases associated with aberrant expression of
XX beta-amyloid peptide, e.g. Alzheimer's disease.
```


PS Claim 4; Page 84-85; 85pp; English.

XX The present invention relates to novel human beta-amyloid peptide (BAP;
CC Abeta, betaAP)-binding (BBP) proteins and polynucleotides encoding such
CC proteins. BBP sequences are useful to diagnose and/or treat diseases
CC associated with aberrant expression of human BAP such as Alzheimer's
CC disease (AD). They are used to generate transgenic animals. Sequences
CC of the invention are also used in gene therapy. The present sequence
CC is human BBP-1 protein.

XX Sequence 269 AA;

Query Match 100.0%; Score 33; DB 24; Length 269;

Best Local Similarity 100.0%; Pred. No. 3; 9e-24;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VAAVSLFLQWLGADREYLGYPALGLKFKCTVG 33

DB 185 VAAVSLFLQWLGADREYLGYPALGLKFKCTVG 217

RESULT 4

AAV12358

ID AAV12358 standard; Protein; 139 AA.

XX

AC AAV12358;

XX

DI 17-JUN-1999 (first entry)

XX

DE Human 5' EST secreted protein SEQ ID NO:389.

XX

KW Human; secreted protein; EST; expressed sequence tag; diagnosis;

KW forensic; gene therapy; chromosome mapping; signal peptide;

KW upstream regulatory sequence; cytokine activity; cell proliferation;

KW differentiation; haematopoiesis regulation; tissue growth regulation;

KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;

KW thrombolytic; anti-inflammatory; tumour inhibition.

XX

OS Homo sapiens.

XX

PN WO9906548-A2.

XX

PD 11-FEB-1999.

XX

PF 31-JUL-1998; 98WO-IB01222.

XX

PR 01-AUG-1997; 97US-0905135.

XX

PA (GEST) GENSET.

XX

PI Duclert A, Dumas Milne Edwards J, Lacroix B;

XX

DR WPI: 1999-153778/13.

XX

DR N-PSDB: AAX41191.

XX

PT New nucleic acids encoding human secreted proteins - obtained from

PT cDNA libraries prepared from e.g. liver, ovary, brain, prostate,

PT kidney, lung, umbilical cord, placenta and colon tissue

XX

PS Claim 27; Page 714-715; 824pp; English.

XX

CC AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for

CC human secreted proteins, and encode the proteins given in AAY12261 to

CC AAY12514, respectively. The proteins given represent the signal peptide

CC and an N-terminal fragment of a secreted protein. The nucleic acid

CC sequences can be used for producing secreted human gene products. They

CC can also be used to develop products for diagnosis and therapy. The

CC proteins obtained may have cytokine activity, cell

CC proliferation/differentiation activity, haematopoiesis regulating

CC activity, tissue growth regulating activity, reproductive hormone

CC regulating activity, chemotactic/chemokinetic activity, haemostatic and

CC thrombolytic activity, receptor/ligand activity, anti-inflammatory

CC activity, tumour inhibition activity or other activities. The products

CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptide can be used for
CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell.

XX Sequence 139 AA;

Query Match 51.5%; Score 17; DB 20; Length 139;

Best Local Similarity 100.0%; Pred. No. 6; 9e-09;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VAAVSLFLQWLGADRF 17

DB 123 VAAVSLFLQWLGADRF 139

RESULT 5

ABR11574

ID ABR11574 standard; peptide; 225 AA.

XX

AC ABR11574;

XX

DI 11-JAN-2002 (first entry)

XX

DF Human secreted protein homologue, SEQ ID NO:1944.

XX

KW Human; cytokine; cell proliferation; cell differentiation; growth factor;

KW haematopoiesis regulation; tissue growth; immunomodulator; activin;

KW inhibitor; chemotaxis; chemokinesis; thrombolysis; oncogenesis;

KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;

KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;

KW chronic inflammatory condition; proliferative retinopathy;

KW atherosclerosis; coronary heart disease; arterial ischaemia;

KW bone disorder; osteoporosis; vascular growth disorder;

KW tissue regeneration; wound healing; infection; immune disorder;

KW cell culture; drug screening; gene therapy; anti-inflammatory;

KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;

KW cytostatic; osteopathic; vasotropic; cardiac; virucide; antibacterial;

KW antifungal; vulnery; antiulcer.

XX

OS Homo sapiens.

XX

PN W0200157188-A2.

XX

PD 09-AUG-2001.

XX

PF 05-FEB-2001; 2001WO-US03806.

XX

PR 03-FEB-2000; 2000US-C496914.

XX

PR 27-APR-2000; 2000US-0560875.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Drmanac RT;

XX

DR WPI: 2001-457740/49.

XX

DR N-PSDB: ABA08818.

XX

PT Human proteins and DNA encoding sequences useful for preventing,

PT treating or ameliorating a medical condition in a mammalian subject

PT e.g. arthritis and cancer -

XX

PS Claim 20; Page 214; 1963pp; English.

XX

CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and

CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The

CC invention also relates to vectors and recombinant host cells comprising

CC nucleotide of the invention, methods of producing the novel polypeptides,

CC antibodies against the polypeptides, methods of detecting the nucleotide,

CC or polypeptides in a sample, and methods of identifying compounds which

CC bind to polypeptides of the invention. Although novel, many of the

CC polypeptides of the invention have homology to known proteins, thereby

CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell-
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a novel human
 CC polypeptide of the invention.

XX Sequence 225 AA;

Query Match 27.3%; Score 9; DB 22; Length 225;
 Best Local Similarity 100.0%; Pred. No. 0.62;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 12 LGADRYFLG 20
 DB 176 LGADRYFLG 184
 |||||

RESULT 6
 ABB77688
 ID ABB77688 standard; peptide: 24 AA.

XX ABB77688;

DT 01-JUL-2002 (first entry)

DE New peptide vector#4.

XX Intracellular delivery; transfection agent; cancer; infectious disease;
 KW peptide vector.

OS Synthetic.

XX WO200210201-A2.

PN 07-FEB-2002.

XX 26-JUL-2001; 2001WO-US23406.

XX 31-JUL-2000; 2000US-221932P.

XX (ACTI-) ACTIVE MOTIF.

PA (CNRS) CENT NAT RECH SCI.

PI Divida G, Morris M, Mery J, Heitz F, Fernandez J, Archdeacon J;

PI Horndorp K;

XX WPI; 2002-329441/36.

PT Transfection agent that comprises a peptide comprising hydrophobic and

PT hydrophilic domain and having amino acid residues of specified length
 PT is useful for a non-covalent association with and transport of a
 PT heterologous compound into a cell

XX Example 2; Page 61; 156pp; English.

XX The invention relates to a transfection agent comprises a peptide of
 XX about 16 - 30 amino acids in length. Peptides of the invention comprise
 XX a hydrophobic domain, a hydrophilic domain, optionally a spacer sequence
 XX between the domains and a functional group conjugated to at least one
 XX terminal of the peptide. Peptides of the invention are useful for a
 XX non-covalent association with and transport of a heterologous compound
 XX into a cell. They are also useful for promoting the cellular
 XX internalisation of at least one member e.g. peptide, proteins,
 XX antibodies, their derivatives and/or conjugates. They may form part of a
 XX pharmaceutical composition to deliver the compound selected from a
 XX diagnostic or therapeutic compound, to treat at least one condition such
 XX as cancer or an infectious disease, or which targets a cancerous cell or
 XX pathogen-infected cell and to deliver a peptide or inhibitor that
 XX disrupts the activity of the enzyme. The agent of the invention has a
 XX transfection efficiency of at least 5% for at least two of the members of
 XX the group of the compounds. The agent has a good delivery efficiency for
 XX a broad spectrum of compounds and cell types, has a low toxicity, are
 XX easy to handle and easy to formulate in conjunction with the many
 XX different compound types that it can deliver. The peptides are serum
 XX sensitive, thus they bind particularly well for systemic and/or localised
 XX in patients. The current sequence represents a new amphipathic peptide
 XX vector of the invention that contains a cationic nuclear localisation
 XX sequence.

XX Sequence 24 AA;

Query Match 24.2%; Score 8; DB 23; Length 24;

Best Local Similarity 100.0%; Pred. No. 0.71;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LFLGLWGA 14
 DB 3 LFLGLWGA 10
 |||||

RESULT 7
 ABB81176

ID ABB81176 standard; peptide: 27 AA.

XX ABB81176;

DT 25-NOV-2002 (first entry)

DE Signal sequence based peptide I MPS peptide.

XX Red blood cell; RBC; virus; vertebrate; therapeutic; tissue imaging;
 KW gp4.; membrane translocation.

XX Human immunodeficiency virus type 1.

OS Ktesus macaque polyoma virus.

XX WO200260416-A1.

XX 08-AUG-2002.

XX 01-FEB-2002; 2002WO-GB00437.

XX 01-FEB-2001; 2001GB-0G02561.

XX 16-FEB-2001; 2001US-269528P.

XX (GENE-) GENDEL LTD.

XX Mchale AP, Craig R;

XX WPI; 2002-643355/69.

PT Delivering agent to target site in vertebrate comprises loading red

PT blood cell with virus or virus-like particle comprising agent,
PI sensitizing cell, introducing cell into vertebrate and applying energy
XX to release virus particle from cell -
XX
PS Disclosure: Page 55; 87pp; English.
XX
CC The invention relates to delivering an agent to a target site in a
CC vertebrate. The method involves (a) loading a red blood cell (RBC) with a
CC virus or a virus-like particle (I) comprising an agent; (b) sensitizing
CC RBC to render it more susceptible to disruption than unsensitized RBC;
CC (c) introducing RBC into a vertebrate; and (d) applying energy to release
CC (I). Steps (a) and (b) may be performed in any order. (II) (RBC loaded
CC with a virus or a virus-like particle comprising a therapeutic agent) is
CC useful for the delivery of a therapeutic agent to a target site in a
CC vertebrate, or in the preparation of a medicament for delivery of a
CC therapeutic agent to a target site in a vertebrate. (II) is also useful
CC for delivering one or more agents to a vertebrate and for treating or
CC preventing a disease. The method is useful for delivering agents such as
CC those useful for imaging of tissues in vivo or ex vivo, preferably for
CC delivering an agent to a subcellular organelle such as nucleus,
CC mitochondria, Golgi or endoplasmic reticulum. The present sequence
CC represents a signal sequence based peptide: WPS peptide, a chimera of
CC the hydrophobic terminal domain of viral gp41 protein and the nuclear
CC localisation signal (NLS) from the SV40 large antigen. This fragment has
CC been found to be active in membrane translocation.
XX
SQ Sequence 27 AA;

Query Match 24.2%; Score 8; DB 23; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LFLGWLGA 14
| | | | | | | |
Db 3 LFLGWLGA 10

RESULT 8
ABG78990
ID ABG78990 standard; Peptide: 27 AA.

XX AC ABG78990;

XX 15-NOV-2002 (first entry)

XX Cell penetrating peptide Signal-peptide II.

XX Cell penetrating peptide; cancer; tumour; melanoma; thymoma;
KW lymphoma; sarcoma; lung cancer; non-Hodgkin's lymphoma; leukaemia;
KW Hodgkin's lymphoma; uterine cancer; cervical cancer; bladder cancer;
KW kidney cancer; adenocarcinoma; breast cancer; prostate cancer;
KW ovarian cancer; pancreatic cancer; vaccine; dendritic cell;
KW tumour infiltrating lymphocyte; TIL; human leukocyte antigen: HLA;
KW cytostatic.

XX Unidentified.

XX WO200264057-A2.

XX 22-AUG-2002.

XX 15-FEB-2002; 2002WO-US05212.

XX 15-FEB-2001; 2001US-268687P.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX Wang R;

XX WPI: 2002-627577/67.

XX Novel composition for treating a disease in an animal, comprises an
PT immune effector cell and cell penetrating peptide associated with an

PT antigen or antibody -
XX
PS Disclosure: Page 11; 61pp; English.
XX
CC The invention relates to a composition (I) comprising an immune effector
CC cell and a cell penetrating peptide (CPP) associated with an antigen or
CC antibody. Also included are (1) a vaccine comprising (I), CPP
CC associated with an antigen, and a pharmaceutically acceptable carrier
CC and (2) preparing a composition for a disease, by providing (I)
CC and CPP associated with an antigen for disease, and introducing the
CC antigen-associated CPP to (I), where antigen enters into the cell.
CC The antigens are, for example, tumour antigen derived epitopes
CC recognised by tumour infiltrating lymphocytes (TIL) of HLA (human
CC leukocyte antigen) class I or II. The composition is useful for enhancing
CC immunity in an animal to a disease, by administering a mature dendritic
CC cell comprising CPP associated with an antigen to disease, to the animal,
CC such that following the administration, animal is protected from disease,
CC where the animal comprises both CD4+ and CD8+ T cells. It is also useful
CC for treating a disease (e.g. cancer, tumour, melanoma, thymoma,
CC lymphoma, sarcoma, lung cancer, non-Hodgkin's lymphoma, leukaemia,
CC Hodgkin's lymphoma, uterine cancer, cervical cancer, bladder cancer,
CC kidney cancer, adenocarcinoma, breast cancer, prostate cancer,
CC ovarian cancer and pancreatic cancer). The animal is further subjected to
CC a cancer treatment including surgery, radiation, chemotherapy or gene
CC therapy. The administration of (I), preferably dendritic cell is prior
CC to, subsequent to or concurrent with, the cancer treatment. The present
CC sequence is cell penetrating peptide of the invention.

XX Sequence 27 AA;

Query Match 24.2%; Score 8; DB 23; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LFLGWLGA 14
| | | | | | | |
Db 3 LFLGWLGA 10

RESULT 9
AAE23685

XX AAE23685 standard; peptide: 27 AA.

XX AC AAE23685;

XX 10-SEP-2002 (first entry)

XX Fluorescently labelled RNA binding peptide #2.

XX RNA binding protein; mRNA quantification; gene expression.

XX Unidentified.

XX WO200227031-A2.

XX 04-APR-2002.

XX 28-SEP-2001; 2001WO-US0438.

XX 28-SEP-2000; 2000US-236407P.

XX (CELL-) CELLONICS INC.

XX Busa WB;

XX WPI: 2002-452280/48.

XX Quantifying target gene expression in living cells that possess a
PT target gene of interest tagged with the binding site for an RNA binding
PT protein and fluorescently labeled RNA binding polypeptide including an
XX RNA binding domain -
XX Claim 45; Page 44; 51pp; English.

XX The present invention relates to a method of quantifying the expression
 CC of target genes in living cells. The method involves providing cells that
 CC possess a target gene of interest which has been tagged with the binding
 CC site for an RNA binding protein and a fluorescently labelled RNA binding
 CC polypeptide that includes an RNA binding domain and calculating the
 CC quantity of target gene expression in the cells using fluorescence
 CC signalling techniques. The method is useful for quantifying expression
 CC of one or more target genes in living cells which comprise two or more
 CC distinct populations of cells. It is used to quantitate the expression
 CC of any target gene, including expression of protein-encoding messenger
 CC RNA genes, ribosomal RNA encoding genes and transfer RNA encoding genes
 CC so long as the RNA expression product from the target gene possesses a
 CC sequence or structure (the RNA tag) that is bound specifically by the
 CC RNA binding polypeptide being used. The present sequence is a
 CC fluorescently labelled RNA binding peptide.

XX Sequence 27 AA;
 SO

Query Match 24.2%; Score 8; DB 23; Length 27;
 Best Local Similarity 100.0%; Pred. No. 0.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LFLGWLGA 14
 Db 11111111
 3 LFLGWLGA 10

RESULT 10
 ABB77687
 ID ABB77687 standard; peptide: 27 AA.
 XX
 AC ABB77687;
 XX
 DT 01-JUL-2002 (first entry)
 XX
 DE New peptide vector#3.
 XX
 KW Intracellular delivery; transfection agent; cancer; infectious disease;
 KW peptide vector.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 7
 FT /note- "residue may be substituted with Phe"
 FT Misc-difference 23
 FT /note- "residue may be substituted with Ser"
 XX
 FN WO200210201-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 26-JUL-2001; 2001WO-US23406.
 XX
 PK 31-JUL-2000; 2000US-221932P.
 XX
 XX (ACT1-) ACTIVE MOTIF.
 PA (CNRS) CENT NAT RECH SCI.
 PA
 PA Divida G, Morris M, Mery J, Heitz F, Fernandez J, Archdeacon J;
 PI Horndorp K;
 XX
 XX WPI; 2002-329441/36.
 DR
 XX
 XX Transfection agent that comprises a peptide comprising hydrophobic and
 PT hydrophilic domain and having amino acid residues of specified length.
 PT is useful for a non-covalent association with and transport of a
 PT heterologous compound into a cell -
 XX Example 2; Page 61; 156pp; English.
 PS
 CC The invention relates to a transfection agent comprises a peptide of

CC about 16 - 30 amino acids in length. Peptides of the invention comprise
 CC a hydrophobic domain, a hydrophilic domain, optionally a spacer sequence
 CC between the domains and a functional group conjugated to at least one
 CC terminal of the peptide. Peptides of the invention are useful for a
 CC non-covalent association with and transport of a heterologous compound
 CC into a cell. They are also useful for promoting the cellular
 CC internalisation of at least one member e.g. peptide, proteins,
 CC antibodies, their derivatives and/or conjugates. They may form part of a
 CC pharmaceutical composition to deliver the compound selected from a
 CC diagnostic or therapeutic compound, to treat at least one condition such
 CC as cancer or an infectious disease, or which targets a cancerous cell or
 CC pathogen-infected cell and to deliver a peptide or inhibitor that
 CC disrupts the activity of the enzyme. The agent of the invention has a
 CC transfection efficiency of at least 5% for at least two of the members of
 CC the group of the compounds. The agent has a good delivery efficiency for
 CC a broad spectrum of compounds and cell types; has a low toxicity, are
 CC easy to handle and easy to formulate in conjunction with the many
 CC different compound types that it can deliver. The peptides are serum
 CC sensitive, thus they bode particularly well for systemic and/or localised
 CC in patients. The current sequence represents a new amphipathic peptide
 CC vector of the invention that contains a cationic nuclear localisation
 CC sequence.

XX Sequence 27 AA;
 SO

Query Match 24.2%; Score 8; DB 23; Length 27;
 Best Local Similarity 100.0%; Pred. No. 0.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LFLGWLGA 14
 Db 11111111
 3 LFLGWLGA 10

RESULT 11
 AA078348
 ID AA078348 standard; Peptide: 27 AA.
 XX
 AC AA078348;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Signal sequence based peptide I.
 XX
 KW Membrane translocation signal; signal sequence based peptide 1;
 KW red blood cell vehicle; polypeptide delivery; viral gp41 protein;
 KW simian virus 40; SV40; large antigen.
 XX
 OS Rhesus macaque polyoma virus.
 OS Synthetic.
 XX
 PN WO200207752-A2.
 XX
 PD 31-JAN-2002.
 XX
 PF 24-JUL-2001; 2001WO-GB03327.
 XX
 XX 24-JUL-2000; 2000WO-GB02848.
 PR 09-AUG-2000; 2000WO-GB03056.
 PR 01-FEB-2001; 2001WO-GB00417.
 PR 16-FEB-2001; 2001US-0785802.
 XX
 XX (GEND-) GENDEL LTD.
 PA
 XX Craig R;
 PI
 XX WPI; 2002-280593/32.
 DR
 XX Preparing a red blood cell vehicle suitable for delivering an agent to
 PT a target site in a vertebrate due to loading the red blood cell with an
 PT agent-membrane translocation sequence -
 XX
 PS Disclosure; Page 44; 135pp; English.

XX The invention describes a method of preparing a red blood cell vehicle
 CC suitable for delivering an agent to a target site in a vertebrate
 CC comprising providing a red blood cell and loading the red blood cell with
 CC an agent-MTS (membrane translocation sequence) conjugate. The red blood
 CC cells produced may be used in the preparation of a medicament for
 CC delivery of an agent to or at a target site and of one or more agents to
 CC a vertebrate. The agent is actively released from the red blood cell
 CC vehicle by application of a stimulus to disrupt the red blood cell
 CC chimeric. This sequence represents signal sequence based peptide 1, a
 CC chimeric peptide of the hydrophobic terminal domain of the viral gp41
 CC protein and the nuclear localisation signal from simian virus 40
 CC (SV40) large antigen, one of the membrane translocation peptides tested
 CC in the invention.

XX Sequence 27 AA;
 CC
 CC Query Match 24.2%; Score 8; DB 23; Length 27;
 CC Best Local Similarity 100.0%; Pred. No. 0.8;
 CC Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LFLGLWGA 14
 Db 3 LFLGLWGA 10

RESULT 12
 ABG75506
 ID ABG75506 standard; Peptide: 27 AA.
 AC
 XX
 XX
 XX 14-APR-2003 (first entry)
 XX
 XX Signal-sequence-based peptide 1.
 XX
 XX Translocation; delivery vehicle; agent-membrane translocation sequence;
 XX MTS; membrane translocation sequence; plasma membrane; red blood cell;
 XX immunisation; antigen; intracellular delivery; therapeutic; HIV-1;
 XX trans-activating protein; Tat; Antennapedia homeodomain protein; Antp-HD;
 XX VP22 protein; HSV-VP22; signal-sequence-based peptide; Transportin; NLS;
 XX Amphiphilic model peptide; SV40; Simian virus 40; gp41;
 XX nuclear localisation sequence.
 XX
 XX Chimeric - Viruses sp.
 XX Chimeric - Rhesus macaque polyoma virus
 XX
 XX US2002151004-A1.
 XX
 XX 17-OCT-2002.
 XX
 XX 16-FEB-2001; 2001US-0785802.
 XX
 XX 24-JUL-2000; 2000GB-0002848.
 XX 09-APR-2000; 2000GB-0003056.
 XX 22-DEC-2000; 2000US-0748063.
 XX 22-DEC-2000; 2000US-0748789.
 XX
 XX (CRAI/) CRAIG R.
 XX
 XX Craig R;
 XX
 XX WPI: 2003-182503/18.
 XX
 XX Preparing delivery vehicles, particularly red blood cells, for the
 XX intracellular delivery of a therapeutic agent to a target site, by
 XX loading a cell with an agent-MTS conjugate, which comprises a membrane
 XX translocation sequence
 XX
 XX Disclosure; Page 8; 43pp; English.

XX The invention discloses a method for preparing a delivery vehicle for
 CC delivering an agent to a target site in a vertebrate. The method

CC comprises loading a cell with an agent-membrane translocation sequence
 CC (MTS) conjugate, which contains a membrane translocation sequence
 CC enabling the agent to cross the plasma membrane of a cell. Also disclosed
 CC is a pharmaceutical composition comprising a red blood cell for delivery
 CC of an agent to a vertebrate, the red blood cell comprising the novel
 CC agent-MTS conjugate and a method of immunisation of an animal with an
 CC antigen. The method is useful for preparing delivery vehicles,
 CC particularly a red blood cell, for the intracellular delivery of a
 CC therapeutic agent to a target site. The method is particularly useful for
 CC enabling an agent to cross the plasma membrane of a target cell, and for
 CC selectively releasing the agent-MTS conjugate at a target site to
 CC facilitate the uptake of the agent by the cells at the target site. The
 CC preferred MTS sequences are from HIV-1 trans-activating protein (Tat),
 CC Drosophila Antennapedia homeodomain protein (Antp-HD), Herpes Simplex-1
 CC virus VP22 protein (HSV-VP22), a signal-sequence-based peptide and a
 CC Transportin and Amphiphilic model peptide. The sequence presented is the
 CC Signal-sequence-based peptide 1, which comprises the hydrophobic terminal
 CC domain of the viral gp41 protein and the nuclear localisation sequence
 CC (NLS) from the SV40 large antigen.

XX Sequence 27 AA;
 CC
 CC Query Match 24.2%; Score 8; DB 24; Length 27;
 CC Best Local Similarity 100.0%; Pred. No. 0.8;
 CC Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LFLGLWGA 14
 Db 3 LFLGLWGA 10

RESULT 13
 ABG82542
 ID ABG82542 standard; peptide; 27 AA.
 XX
 XX AC ABG82542;
 XX
 XX 04-FEB-2003 (first entry)
 XX
 XX Signal-sequence based peptide 1 MPS peptide.
 XX
 XX Red blood cell; transgenic; cytostatic; immunosuppressive; SV40; MPS;
 XX anti-rheumatic; antiarthritic; dermatological; antithyroid; nootropic;
 XX neuroprotective; cerebroprotective; cardiac; antiarrhythmic; litholytic;
 XX antiinflammatory; nephrotropic; antidiabetic; analgesic; osteopathic; gp41.
 XX
 XX Synthetic.
 XX Rhesus macaque polyoma virus.
 XX
 XX WO200257436-A2.
 XX
 XX 25-JUL-2002.
 XX
 XX 16-JAN-2002; 2002WO-GH00169.
 XX
 XX 19-JAN-2001; 2001GB-0001469.
 XX 23-JAN-2001; 2001US-264808P.
 XX 01-FEB-2001; 2001WO-GH00417.
 XX 16-FEB-2001; 2001US-0785802.
 XX 07-MAR-2001; 2001GB-0005631.
 XX 29-MAR-2001; 2001US-279803P.
 XX 24-JUL-2001; 2001WO-GH03327.
 XX
 XX (GEND-) GENDEL LTD.
 XX
 XX McHale AP, Craig RK;
 XX
 XX WPI: 2003-040499/03.
 XX
 XX Producing red blood cells for polypeptide delivery, by obtaining a red
 XX blood cell containing the polypeptide from a transgenic animal
 XX expressing a gene encoding the polypeptide, and sensitizing the cell

CC The present sequence is the beta-amyloid peptide (BAP) binding protein-3
CC (BBP3). It is an integral membrane protein, that traverse the membrane
CC twice. It is related to G protein-coupled receptor (GPCR) protein
CC superfamily. It interacts with G-alpha proteins and regulates the
CC activity of G-protein signalling pathways. BBP genes are widely expressed
CC in neuronal cells of nonhuman primate (NHP) brain and overexpressed in
CC some tumours. It functions as a suppressor of apoptosis induction. BBP
CC proteins are used as immunogens to raise antibodies, useful as
CC therapeutics and as antigens in solid phase assays. They are also useful
CC as reagents to identify molecules which effect the interaction of BBP and
CC a cloned protein, that are useful in the treatment or prevention of
CC diseases associated with apoptosis. The polynucleotides are useful for
CC diagnostics.
xx

SQ Sequence 221 AA:

Query Match 24.2%; Score 8; DB 21; Length 221;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 GADREYLG 20
Db 174 GADREYLG 181
|||||||

Search completed: September 26, 2003, 17:39:25
Job time : 17.7106 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2003, 17:36:55 : Search time 5.89787 Seconds
(without alignments)
236.739 Million cell updates/sec

Title: US-09-852-100A-2_COPY_185_217

Perfect score: 33

Sequence: 1 VAVLSLFLGW:GNDRFLGYLPALGLLK:CTVG 33

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents, NA:*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/iaa/6C_COMB.pep:*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	24.2	358	4	US-09-411-628-11
2	7	21.2	451	4	US-09-345-473E-45
3	7	21.2	462	4	US-09-328-352-6889
4	7	21.2	483	4	US-09-252-991A-29267
5	7	21.2	617	4	US-09-345-473E-44
6	6	18.2	46	4	US-09-489-847-131
7	6	18.2	70	4	US-09-205-258-298
8	6	18.2	72	3	US-09-905-223-314
9	6	18.2	101	4	US-09-489-847-306
10	6	18.2	117	4	US-09-228-986-111
11	6	18.2	123	3	US-08-675-508-2
12	6	18.2	123	3	US-09-203-939-2
13	6	18.2	123	3	US-09-203-939-6
14	6	18.2	123	3	US-09-251-835-2
15	6	18.2	123	3	US-09-251-835-6
16	6	18.2	123	3	US-09-318-503-2
17	6	18.2	123	3	US-09-318-503-6
18	6	18.2	123	3	US-09-038-261A-2
19	6	18.2	123	3	US-09-038-261A-6
20	6	18.2	123	4	US-09-564-329A-2
21	6	18.2	123	4	US-09-564-329A-6
22	6	18.2	209	2	US-08-808-550-31
23	6	18.2	261	4	US-09-252-991A-17349
24	6	18.2	281	4	US-09-252-991A-33042
25	6	18.2	325	4	US-09-252-991A-16657
26	6	18.2	349	4	US-09-162-524-3
27	6	18.2	350	4	US-08-637-670-25

```

28      6  18.2      360  4  US-09-328-352-5126      Sequence 5126, Ap
29      6  18.2      380  4  US-09-252-991A-24774      Sequence 24774, A
30      6  18.2      384  1  US-08-221-750A-11      Sequence 11, Appl
31      6  18.2      389  4  US-09-252-991A-21202      Sequence 21202, A
32      6  18.2      396  4  US-09-198-452A-147      Sequence 147, Ap
33      6  18.2      404  4  US-09-328-352-6747      Sequence 6747, Ap
34      6  18.2      421  4  US-09-252-991A-30177      Sequence 30177, A
35      6  18.2      430  4  US-09-252-991A-31388      Sequence 31388, A
36      6  18.2      432  1  US-08-476-008-61      Sequence 61, Appl
37      6  18.2      432  1  US-08-306-363-61      Sequence 61, Appl
38      6  18.2      432  1  US-08-833-485-61      Sequence 61, Appl
39      6  18.2      432  3  US-09-137-440-61      Sequence 61, Appl
40      6  18.2      434  1  US-08-162-392-4      Sequence 4, Appl
41      6  18.2      434  1  US-08-643-300-4      Sequence 4, Appl
42      6  18.2      434  1  US-08-643-297-4      Sequence 4, Appl
43      6  18.2      434  1  US-08-643-298-4      Sequence 4, Appl
44      6  18.2      434  2  US-08-643-301-4      Sequence 4, Appl
45      6  18.2      434  5  PCT-US94-14095-4      Sequence 4, Appl

```

ALIGNMENTS

```

RESULT 1
US-09-411-628-11
: Sequence 11, Application US/09411628
: Patent No. 6428994
: GENERAL INFORMATION:
: APPLICANT: University of Southern California
: TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
: FILE REFERENCE: 13761-707
: CURRENT APPLICATION NUMBER: US/09/411,628
: CURRENT FILING DATE: 1999-10-01
: EARLIER APPLICATION NUMBER: US 60/102,906
: EARLIER FILING DATE: 1998-10-02
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 11
: LENGTH: 358
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-411-628-11

```

```

Query Match: 24.2% Score 8; DB 4; Length 358;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 8; Conservative C; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      21 YPALGLLK 28
      |||||
Pb      259 YPALGLLK 265

```

```

RESULT 2
US-09-345-473E-45
: Sequence 45, Application US/09345473E
: Patent No. 6558903
: GENERAL INFORMATION:
: APPLICANT: Hodge, Martin
: TITLE OF INVENTION: No. 6558903el Kinases and Uses Thereof
: FILE REFERENCE: 35800/183781
: CURRENT APPLICATION NUMBER: US/09/345,473E
: CURRENT FILING DATE: 1999-06-30
: NUMBER OF SEQ ID NOS: 62
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 45
: LENGTH: 451
: TYPE: PRT
: ORGANISM: Mus musculus
US-09-345-473E-45

```

```

Query Match: 21.2% Score 7; DB 4; Length 451;
Best Local Similarity 100.0%; Pred. No. 21;

```



```
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 ALSFLG 10
Db 416 ALSFLG 422

RESULT 3
US-09-328-352-6888
: Sequence 6888, Application US/09328352
: Patent No. 6562958
: GENERAL INFORMATION:
: APPLICANT: Gary L. Brceton et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
: FILE REFERENCE: GTG99-03PA
: CURRENT APPLICATION NUMBER: US/09/328,352
: CURRENT FILING DATE: 1999-06-04
: NUMBER OF SEQ ID NOS: 8252
: SEQ ID NO 5888
: LENGTH: 462
: TYPE: PRT
: ORGANISM: Acinetobacter baumannii
US-09-328-352-6888

Query Match 21.2%; Score 7; DB 4; Length 462;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 ALSFLG 10
Db 334 ALSFLG 340

RESULT 4
US-09-252-991A-29267
: Sequence 29267, Application US/09252991A
: Patent No. 6551795
: GENERAL INFORMATION:
: APPLICANT: Marc J. Rubenfield et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
: FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252,991A
: CURRENT FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/074,788
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
: PRIOR FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 33142
: SEQ ID NO 29267
: LENGTH: 483
: TYPE: PRT
: ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29267

Query Match 21.2%; Score 7; DB 4; Length 483;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 22 PALGLLK 28
Db 438 PALGLLK 444

RESULT 5
US-09-345-473E-44
: Sequence 44, Application US/09345473E
: Patent No. 6558903
: GENERAL INFORMATION:
: APPLICANT: Hodges, Martin
: TITLE OF INVENTION: No. 6558903ei Kinases and Uses Thereof
: FILE REFERENCE: 35800/183781
```

```
: CURRENT APPLICATION NUMBER: US/09/345,473E
: CURRENT FILING DATE: 1999-06-30
: NUMBER OF SEQ ID NOS: 62
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 44
: LENGTH: 617
: TYPE: PRT
: ORGANISM: Mus musculus
US-09-345-473E-44

Query Match 21.2%; Score 7; DB 4; Length 617;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 ALSFLG 10
Db 582 ALSFLG 588

RESULT 6
US-09-489-847-131
: Sequence 131, Application US/09489847
: Patent No. 6475195
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: 98 Human Secreted Proteins
: FILE REFERENCE: P2031E1
: CURRENT APPLICATION NUMBER: US/09/489,847
: CURRENT FILING DATE: 2000-01-24
: EARLIER APPLICATION NUMBER: PCT/US99/17130
: EARLIER FILING DATE: 1999-07-29
: EARLIER APPLICATION NUMBER: 60/094,657
: EARLIER FILING DATE: 1998-07-30
: EARLIER APPLICATION NUMBER: 60/095,486
: EARLIER FILING DATE: 1998-08-05
: EARLIER APPLICATION NUMBER: 60/096,319
: EARLIER FILING DATE: 1998-08-12
: EARLIER APPLICATION NUMBER: 60/095,454
: EARLIER FILING DATE: 1998-08-06
: EARLIER APPLICATION NUMBER: 60/095,455
: EARLIER FILING DATE: 1998-08-06
: NUMBER OF SEQ ID NOS: 376
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 131
: LENGTH: 46
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (46)
: OTHER INFORMATION: Xaa equals stop translation
US-09-489-847-131

Query Match 18.2%; Score 6; DB 4; Length 46;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 24 LGLLKF 29
Db 10 LGLLKF 15

RESULT 7
US-09-205-258-298
: Sequence 298, Application US/09205258
: Patent No. 6525174
: GENERAL INFORMATION:
: APPLICANT: Young et al.
: TITLE OF INVENTION: 207 Human Secreted Proteins
: FILE REFERENCE: P2007P1
: CURRENT APPLICATION NUMBER: US/09/205,258
: CURRENT FILING DATE: 1998-12-04
: EARLIER APPLICATION NUMBER: PCT/US98/11422
```

EARLIER FILING DATE: 1998-06-04
EARLIER APPLICATION NUMBER: 60/048,885
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,375
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,881
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,880
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,896
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,020
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,876
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,884
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,894
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,899
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO: 298
LENGTH: 70
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (70)
OTHER INFORMATION: Xaa equals stop translation
US-09-205-258-298
Query Match 18.2%, Score 6; DB 4; Length 70;
Best Local Similarity 100.0%, Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 4 ALSLFL 9
IIIIII
CO 4 ALSLFL 9
RESULT 8
US-08-305-223-314
Sequence 314, Application US/08905223
Patent No. 6222029
GENERAL INFORMATION:
APPLICANT: Edwards, Jean-Baptiste D.
APPLICANT: Duclert, Aymeric
APPLICANT: Lacroix, Bruno
TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
NUMBER OF SEQUENCES: 503
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 501 West Broadway
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-3505
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPIER: IBM PC compatible
OPERATING SYSTEM: Win95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/905,223
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 314:
SEQUENCE CHARACTERISTICS:
LENGTH: 72 amino acids
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
ORIGINAL SOURCE:
ORGANISM: Homo Sapiens
TISSUE TYPE: Brain
FEATURE:
NAME/KEY: sig_peptide
LOCATION: -42...-1
IDENTIFICATION METHOD: Von Heijne matrix

OTHER INFORMATION: score 6.1
OTHER INFORMATION: seq DLAVALSLPFWT/ES
US-08-905-223-314

Query Match 18.2%, Score 6; DB 3; Length 72;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AVALSL 7
DB 32 AVALSL 37

RESULT 9
US-09-489-847-306
; Sequence 306, Application US/09499847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P203191
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCI/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 50/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 306
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-489-847-306

Query Match 18.2%, Score 6; DB 4; Length 101;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AVALSL 7
DB 89 AVALSL 94

RESULT 10
US-09-228-986-111
; Sequence 111, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niele
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 111
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-228-986-111

Query Match 18.2%, Score 6; DB 4; Length 117;

Best Local Similarity 100.0%; Pred. No. 60;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VAVALS 6
DB 37 VAVALS 42

RESULT 11
US-06-675-508-2
; Sequence 2, Application US/08675508
; Patent No. 5856136
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: NOVEL HUMAN STEM CELL ANTIGENS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,508
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0066 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 123 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: SCAR-2
; CLONE:
US-08-675-508-2

Query Match 18.2%, Score 6; DB 2; Length 123;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 PALGLL 27
DB 111 PALGLL 116

RESULT 12
US-09-203-939-2
; Sequence 2, Application US/09203939
; Patent No. 6258939
; GENERAL INFORMATION:
; APPLICANT: Reiter, Robert E.
; APPLICANT: Witte, Owen N.
; TITLE OF INVENTION: PSMA: PROSTATE STEM CELL ANTIGEN AND USES THEREOF
; FILE REFERENCE: 30435.54US11
; CURRENT APPLICATION NUMBER: US/09/203,939
; CURRENT FILING DATE: 2000-12-02
; PRIOR APPLICATION NUMBER: 08/814,279
; PRIOR FILING DATE: 1997-03-10
; PRIOR APPLICATION NUMBER: 60/071,141

```

; PRIOR FILING DATE: 1998-01-12
; CURRENT APPLICATION NUMBER: 60/074,675
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 09/038,261
; PRIOR FILING DATE: 1998-03-10
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 123
; TYPE: PRT
; ORGANISM: HUMAN PSCA (hPSCA)
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (50)..(64)
; NAME/KEY: SITE
; LOCATION: (71)..(82)
; NAME/KEY: SITE
; LOCATION: (67)..(81)
; US-09-203-939-2

```

```

Query Match      18.2%; Score 6; DB 3; Length 123;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      22 PALGGL 27
DB      111 PALGGL 116

```

RESULT 13

```

US-09-203-939-5
; Sequence 6, Application US/09201939
; Patent No. 6258939
; GENERAL INFORMATION:
; APPLICANT: Reiter, Robert E.
; APPLICANT: Witte, Owen N.
; TITLE OF INVENTION: PSCA: PROSTATE STEM CELL ANTIGEN AND USES THEREOF
; FILE REFERENCE: 30435.54US1
; CURRENT APPLICATION NUMBER: US/09/203,939
; PRIOR FILING DATE: 2000-12-02
; PRIOR FILING DATE: 08/314,279
; PRIOR FILING DATE: 1997-03-10
; PRIOR APPLICATION NUMBER: 60/071,141
; PRIOR FILING DATE: 1998-01-12
; PRIOR APPLICATION NUMBER: 60/074,675
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 09/038,261
; PRIOR FILING DATE: 1998-03-10
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 123
; TYPE: PRT
; ORGANISM: HUMAN PSCA (hPSCA)
; US-09-203-939-6

```

```

Query Match      18.2%; Score 6; DB 3; Length 123;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      22 PALGGL 27
DB      111 PALGGL 116

```

RESULT 14

```

US-09-251-835-2
; Sequence 2, Application US/09251835A
; Patent No. 6261789
; GENERAL INFORMATION:
; APPLICANT: Reiter, Robert E.
; APPLICANT: Witte, Owen N.
; TITLE OF INVENTION: PSCA: PROSTATE STEM CELL ANTIGEN

```

```

; FILE REFERENCE: 30435.54US12
; CURRENT APPLICATION NUMBER: US/09/251,835A
; CURRENT FILING DATE: 1999-02-17
; PRIOR APPLICATION NUMBER: 08/814,279
; PRIOR FILING DATE: 1997-03-10
; PRIOR APPLICATION NUMBER: 60/071,141
; PRIOR FILING DATE: 1998-01-12
; PRIOR APPLICATION NUMBER: 60/074,675
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 09/038,261
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 09/203,939
; PRIOR FILING DATE: 1998-12-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 123
; TYPE: PRT
; ORGANISM: HUMAN PSCA (hPSCA)
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (50)..(64)
; NAME/KEY: SITE
; LOCATION: (71)..(82)
; NAME/KEY: SITE
; LOCATION: (67)..(81)
; US-09-251-835-2

```

```

Query Match      18.2%; Score 6; DB 3; Length 123;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      22 PALGGL 27
DB      111 PALGGL 116

```

RESULT 15

```

US-09-251-835-6
; Sequence 6, Application US/09251835A
; Patent No. 6261789
; GENERAL INFORMATION:
; APPLICANT: Reiter, Robert E.
; APPLICANT: Witte, Owen N.
; TITLE OF INVENTION: PSCA: PROSTATE STEM CELL ANTIGEN
; FILE REFERENCE: 30435.54US12
; CURRENT APPLICATION NUMBER: US/09/251,835A
; CURRENT FILING DATE: 1999-02-17
; PRIOR APPLICATION NUMBER: 08/814,279
; PRIOR FILING DATE: 1997-03-10
; PRIOR APPLICATION NUMBER: 60/071,141
; PRIOR FILING DATE: 1998-01-12
; PRIOR APPLICATION NUMBER: 60/074,675
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 09/038,261
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 09/203,939
; PRIOR FILING DATE: 1998-12-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 123
; TYPE: PRT
; ORGANISM: HUMAN PSCA (hPSCA)
; US-09-251-835-6

```

```

Query Match      18.2%; Score 6; DB 3; Length 123;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      22 PALGGL 27
DB      111 PALGGL 116

```

Search completed: September 26, 2003, 17:43:24
Job time : 6.89787 secs

GenCore version 5.1.6
Copyright. (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2003, 17:41:48 : Search time 11.5149 seconds
(without alignments)
433.624 Million cell updates/sec

Title: US-09-852-100A-2_COPY_185_217
Perfect score: 33
Sequence: 1 VAVALSLFLGWLGLADRFYLGYPALGLLKFTCTVG 33

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 566894 seqs, 151307093 residues

Word size : 0

Total number of hits satisfying chosen parameters: 566894

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PublishedApplications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	100.0	269	9	US-09-852-100A-2
2	33	100.0	269	10	US-09-833-503A-2
3	33	100.0	269	15	US-10-199-881-2
4	8	24.2	21	15	US-10-226-956-288
5	8	24.2	21	15	US-10-211-088-306
6	8	24.2	24	11	US-09-015-914B-28
7	8	24.2	27	10	US-09-785-802A-8
8	8	24.2	27	11	US-09-965-876A-2
9	8	24.2	27	11	US-09-915-914B-27
10	8	24.2	27	15	US-10-077-555-8
11	8	24.2	221	11	US-09-833-503A-6
12	8	24.2	221	11	US-09-992-600A-82
13	8	24.2	221	11	US-09-924-340-82
14	8	24.2	221	11	US-09-974-879-230
15	8	24.2	221	12	US-09-992-095B-82

16	8	24.2	221	12	US-10-154-678-82	Sequence 82, Appl
17	8	24.2	221	12	US-09-999-570-82	Sequence 82, Appl
18	8	24.2	221	15	US-10-060-489-82	Sequence 82, Appl
19	8	24.2	221	15	US-10-000-986-62	Sequence 82, Appl
20	8	24.2	221	15	US-10-199-881-6	Sequence 6, Appl
21	8	24.2	222	11	US-09-305-736-230	Sequence 230, Appl
22	8	24.2	230	11	US-09-796-753-50	Sequence 50, Appl
23	8	24.2	247	11	US-09-796-753-48	Sequence 48, Appl
24	8	24.2	358	12	US-10-174-794-11	Sequence 11, Appl
25	7	21.2	122	9	US-09-867-550-80	Sequence 80, Appl
26	7	21.2	173	11	US-09-946-374-332	Sequence 332, Appl
27	7	21.2	173	12	US-10-015-387A-332	Sequence 332, Appl
28	7	21.2	173	12	US-10-006-130A-332	Sequence 332, Appl
29	7	21.2	173	12	US-10-006-172A-332	Sequence 332, Appl
30	7	21.2	173	12	US-10-015-392A-332	Sequence 332, Appl
31	7	21.2	173	12	US-10-017-253A-332	Sequence 332, Appl
32	7	21.2	173	12	US-10-017-306A-332	Sequence 332, Appl
33	7	21.2	173	15	US-10-006-856A-332	Sequence 332, Appl
34	7	21.2	173	15	US-10-006-818A-332	Sequence 332, Appl
35	7	21.2	173	15	US-10-015-393A-332	Sequence 332, Appl
36	7	21.2	173	15	US-10-015-869A-332	Sequence 332, Appl
37	7	21.2	173	15	US-10-012-121A-332	Sequence 332, Appl
38	7	21.2	173	15	US-10-006-116A-332	Sequence 332, Appl
39	7	21.2	173	15	US-10-006-117A-332	Sequence 332, Appl
40	7	21.2	173	15	US-10-017-527A-332	Sequence 332, Appl
41	7	21.2	173	15	US-10-013-913A-332	Sequence 332, Appl
42	7	21.2	173	15	US-10-007-194A-332	Sequence 332, Appl
43	7	21.2	173	15	US-10-013-430A-332	Sequence 332, Appl
44	7	21.2	173	15	US-10-011-671A-332	Sequence 332, Appl
45	7	21.2	173	15	US-10-012-755A-332	Sequence 332, Appl

ALIGNMENTS

RESULT :
US-09-852-100A-2
: Sequence 2, Application US/09852100A
: Patent No. US20020058267A1
: GENERAL INFORMATION:
: APPLICANT: American Home Products
: TITLE OF INVENTION: Beta-amyloid Peptide-Binding Proteins and Polynucleotides Enc
: FILE OF INVENTION: Same
: FILE REFERENCE: AHP981261P2
: CURRENT APPLICATION NUMBER: US/09/852,100A
: CURRENT FILING DATE: 2001-05-09
: PRIOR APPLICATION NUMBER: US 09/172,990
: PRIOR FILING DATE: 1998-10-14
: PRIOR APPLICATION NUMBER: US 60/104,104
: PRIOR FILING DATE: 1998-10-13
: PRIOR APPLICATION NUMBER: PTC/US99/21621
: PRIOR FILING DATE: 1999-10-13
: PRIOR APPLICATION NUMBER: US 09/060,609
: PRIOR FILING DATE: 1998-04-15
: PRIOR APPLICATION NUMBER: US 60/054,583
: PRIOR FILING DATE: 1997-04-16
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 2
: LENGTH: 269
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-852-100A-2

Query Match 100.0%; Score 33; DB 9; Length 269;
Best Local Similarity 100.0%; Pred. No. 9.4e-24;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VAVALSLFLGWLGLADRFYLGYPALGLLKFTCTVG 33
|||||
DB 185 VAVALSLFLGWLGLADRFYLGYPALGLLKFTCTVG 217

```
RESULT 2
US-09-833-503A-2
; Sequence 2, Application US/09833503A
; Patent No. US20020146760A1
; GENERAL INFORMATION:
; APPLICANT: Ozenberger, Bradley A
; APPLICANT: Kajkowski, Eileen M
; APPLICANT: Lo, Ching-Hsiung F
; APPLICANT: American Home Products Corporation
; TITLE OF INVENTION: No. US20020146760A1 G-Protein-Coupled Receptor-Like Proteins and Methods of Using
; TITLE OF INVENTION: Polynucleotides Encoded By Them, and Methods of Using
; TITLE OF INVENTION: Same
; FILE REFERENCE: AHP98155-00PCT
; CURRENT APPLICATION NUMBER: US/09/833,503A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/104,104
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-503A-2

Query Match          100.0%; Score 33; DB 10; Length 269;
Best Local Similarity 100.0%; Pred. No. 9,4e-24;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VAVALSFLGWLGNADRYLGYPALGLLKFTCTVG 33
      |||
DB 185 VAVALSFLGWLGNADRYLGYPALGLLKFTCTVG 217

RESULT 3
US-10-199-881-2
; Sequence 2, Application US/10199881
; Publication No. US20030096356A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: No. US20030096356A1 G-Protein-Coupled Receptor-Like Proteins and Methods of Using Same
; FILE REFERENCE: AHP98155C1
; CURRENT APPLICATION NUMBER: US/10/199,881
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/US95/21621
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: US 90/833,5081
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US 60/104,104
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 2
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-881-2

Query Match          100.0%; Score 33; DB 15; Length 269;
Best Local Similarity 100.0%; Pred. No. 9,4e-24;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VAVALSFLGWLGNADRYLGYPALGLLKFTCTVG 33
      |||
DB 185 VAVALSFLGWLGNADRYLGYPALGLLKFTCTVG 217

RESULT 4
US-10-226-956-288
; Sequence 288, Application US/10226955
; Publication No. US20030060399A1
; GENERAL INFORMATION:
; APPLICANT: Divita, Gilles
; APPLICANT: Fernandez, Joseph
; APPLICANT: Heitz, Frederic
```

```
; APPLICANT: Brophy, Colleen
; APPLICANT: Komalavilas, Padmini
; APPLICANT: Panitch, Alyssa
; APPLICANT: Joshi, Lokesh
; APPLICANT: Seal, Brandon L.
; TITLE OF INVENTION: REAGENTS AND METHODS FOR SMOOTH MUSCLE THERAPIES
; FILE REFERENCE: ASU-106-US
; CURRENT APPLICATION NUMBER: US/10/226,956
; CURRENT FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: 60/314,535
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 320
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 288
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-10-226-956-288

Query Match          24.2%; Score 8; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LFLGWLGA 14
      |||
DB 3 LFLGWLGA 10

RESULT 5
US-10-211-088-306
; Sequence 306, Application US/10211088
; Publication No. US20030104479A1
; GENERAL INFORMATION:
; APPLICANT: Bright, Gary R.
; APPLICANT: Premkumar, D. David
; APPLICANT: Chen, Yih-Tai
; TITLE OF INVENTION: No. US20030104479A1 Fusion Proteins And Assays For Molecular
; FILE REFERENCE: 01-1022-US
; CURRENT APPLICATION NUMBER: US/10/211,088
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/309,395
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/341,589
; PRIOR FILING DATE: 2001-12-13
; NUMBER OF SEQ ID NOS: 366
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 306
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Protein-derived transport peptide
US-10-211-088-306

Query Match          24.2%; Score 8; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LFLGWLGA 14
      |||
DB 3 LFLGWLGA 10

RESULT 6
US-09-915-914B-28
; Sequence 28, Application US/09915914B
; Publication No. US20030119725A1
; GENERAL INFORMATION:
; APPLICANT: Divita, Gilles
; APPLICANT: Fernandez, Joseph
; APPLICANT: Heitz, Frederic
```

```
; APPLICANT: Morris, May
; APPLICANT: Mery, Jean
; APPLICANT: Archdeacon, John
; APPLICANT: Horndorp, Kyle
; TITLE OF INVENTION: PEPTIDE-MEDIATED DELIVERY OF MOLECULES INTO CELLS
; FILE REFERENCE: AM-00105.P.1.1
; CURRENT APPLICATION NUMBER: US/09/915,914B
; CURRENT FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: US 60/221,932
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic sequence
US-09-915-914B-28

Query Match      24.2%; Score 8; DB 11; Length 24;
Best Local Similarity 100.0%; Pred. No. C.56;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      7 LFLGWLGA 14
Db      3 LFLGWLGA 10

RESULT 7
US-09-785-802A-8
; Sequence 8, Application US/09785802A
; Patent No. US20020151004A1
; GENERAL INFORMATION:
; APPLICANT: Craig, Roger
; TITLE OF INVENTION: DELIVERY VEHICLES AND METHODS FOR USING THE SAME
; FILE REFERENCE: 11067/2035
; CURRENT APPLICATION NUMBER: US/09/785,802A
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 09/748,06
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/748,789
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-785-802A-8

Query Match      24.2%; Score 8; DB 10; Length 27;
Best Local Similarity 100.0%; Pred. No. C.63;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      7 LFLGWLGA 14
Db      3 LFLGWLGA 10

RESULT 8
US-09-965-876A-2
; Sequence 2, Application US/09965876A
; Publication No. US20030096243A1
; GENERAL INFORMATION:
; APPLICANT: Cellomics, Inc.
; APPLICANT: Busa, William B
; TITLE OF INVENTION: Methods and Reagents for Live-cell Gene Expression Quantification
; FILE REFERENCE: 00-789-A
; CURRENT APPLICATION NUMBER: US/09/965,876A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/236,407
; PRIOR FILING DATE: 2000-09-28

; APPLICANT: Divita, Gilles
; APPLICANT: Fernandez, Joseph
; APPLICANT: Heitz, Frederic
; APPLICANT: Morris, May
; APPLICANT: Mery, Jean
; APPLICANT: Archdeacon, John
; APPLICANT: Horndorp, Kyle
; TITLE OF INVENTION: PEPTIDE-MEDIATED DELIVERY OF MOLECULES INTO CELLS
; FILE REFERENCE: AM-00105.P.1.1
; CURRENT APPLICATION NUMBER: US/09/915,914B
; CURRENT FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: US 60/221,932
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic sequence
US-09-915-914B-27

Query Match      24.2%; Score 8; DB 11; Length 27;
Best Local Similarity 100.0%; Pred. No. C.63;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      7 LFLGWLGA 14
Db      3 LFLGWLGA 10

RESULT 9
US-09-915-914B-27
; Sequence 27, Application US/09915914B
; Publication No. US20030119725A1
; GENERAL INFORMATION:
; APPLICANT: Divita, Gilles
; APPLICANT: Fernandez, Joseph
; APPLICANT: Heitz, Frederic
; APPLICANT: Morris, May
; APPLICANT: Mery, Jean
; APPLICANT: Archdeacon, John
; APPLICANT: Horndorp, Kyle
; TITLE OF INVENTION: PEPTIDE-MEDIATED DELIVERY OF MOLECULES INTO CELLS
; FILE REFERENCE: AM-00105.P.1.1
; CURRENT APPLICATION NUMBER: US/09/915,914B
; CURRENT FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: US 60/221,932
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic sequence
US-09-915-914B-27

Query Match      24.2%; Score 8; DB 11; Length 27;
Best Local Similarity 100.0%; Pred. No. C.63;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      7 LFLGWLGA 14
Db      3 LFLGWLGA 10

RESULT 10
US-10-077-555-8
; Sequence 8, Application US/10077555
; Publication No. US20030077289A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Rong-fu
; TITLE OF INVENTION: Use of Cell-Penetrating Peptides to Generate Antitumor Immun;
; FILE REFERENCE: P02373US1/10200806
; CURRENT APPLICATION NUMBER: US/10/077,555
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/268,687
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 27
; TYPE: PRT
```



```

; PRIOR APPLICATION NUMBER: PCT/US98/234355
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: US 60/064,911
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,912
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,983
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,900
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,986
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,987
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,908
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,984
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,985
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/066,094
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,100
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,089
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,095
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,090
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 611
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 230
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (184)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-974-879-230

```

```

Query Match      24.2%; Score 8; DB 11; Length 221;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      13 GADRFYLG 20
Db      174 GADRFYLG 181

```

```

RESULT 15
US-09-992-095B-82
; Sequence 82, Application US/0992095B
; Publication No. US20030157485A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US5.DIV
; CURRENT APPLICATION NUMBER: US/09/992.095B
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574

```

```

; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: Spatent
; SEQ ID NO 82
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: 1..32
US-09-992-095B-82

```

```

Query Match      24.2%; Score 8; DB 12; Length 221;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qy      13 GADRFYLG 20
Db      174 GADRFYLG 181

```

```

Search completed: September 26, 2003, 17:55:13
Job time : 12.5149 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2003, 17:28:28 ; Search time 4.93192 Seconds
(without alignments)
394.688 Million cell updates/sec

Title: us-09-852-100a-2_COPY_185_217

Perfect score: 33

Sequence: 1 VVALSLFLGWLGLADRFYLGYPALGLLKFTCTVG 33

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	24.2	338	1	Q00532 homo sapien
2	7	21.2	215	2	P42978 bacillus s
3	7	21.2	224	1	P56750 homo sapien
4	7	21.2	419	1	P81376 rhodospirillum rubrum
5	7	21.2	489	1	Q9512 homo sapien
6	7	21.2	510	1	P05819 escherichia coli
7	7	21.2	575	1	Q10284 schizosaccharomyces pombe
8	7	21.2	638	1	P53670 rattus norvegicus
9	7	21.2	1039	1	P39180 escherichia coli
10	6	18.2	39	1	P20120 pisum sativum
11	6	18.2	58	1	Q9B257 homo sapien
12	6	18.2	111	1	Q9K819 bacillus thuringiensis
13	6	18.2	123	1	Q43553 homo sapien
14	6	18.2	139	1	Q91121 morone saxatilis
15	6	18.2	160	1	Q95757 arabidopsis thaliana
16	6	18.2	163	1	C8K9T5 buchnera americana
17	6	18.2	167	1	P12357 spinacia oleracea
18	6	18.2	171	1	Q82598 rhizobium meliloti
19	6	18.2	172	1	Q8YEC2 bruceella maritima
20	6	18.2	176	1	Q28553 archaeoglobus fulgidus
21	6	18.2	184	1	Q9K193 vibrio cholerae
22	6	18.2	189	1	TBP1THECE
23	6	18.2	202	1	Q97201 neurospora crassa
24	6	18.2	205	1	G10354 orgyia pseudomaculata
25	6	18.2	206	1	SS72-YEAST
26	6	18.2	207	1	IOLB-ATPAC
27	6	18.2	219	1	END3-BACSU
28	6	18.2	239	1	FLIP-AQUAE
29	6	18.2	239	1	YVIC-BACSU
30	6	18.2	240	1	MNTB-LISIN
31	6	18.2	240	1	MNTB-LISIN
32	6	18.2	242	1	NIFY-AZOVI
33	6	18.2	253	1	PECE-ANASP

RESULT 1

KKIA_HUMAN

AC Q00532; STANDARD: PRT: 358 AA.
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-OCT-2001 (Rel. 40, Last annotation update)
 DE Serine/threonine-protein kinase KIALRE (EC 2.7.1.1) (Cyclin-dependent kinase-like 1).
 GN CDKL1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=92347325; PubMed=1639063;
 RX MeyerSON M., Enders G.H., Wu C.-L., Su L.-K., Gorka C., Nelson C., Harlow E., Tsai L.-H.;
 RT "A family of human cdc2-related protein kinases.";
 RL EMBO J 11:2909-2917(1992).
 SC 1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC CDC2/CDKX SUBFAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL: X66358; CAA47002.1; JOINED.
 FTR: S23383; S23383.
 SSF: P24941; LHCL.
 Genes: HGNC:1781; CDKL1.
 GK: Q00532; .
 MM: 603441; .
 GO: GO:0004693; E: cyclin-dependent protein kinase activity; TAS.
 GO: GO:0006468; P: protein amino acid phosphorylation; TAS.
 GO: GO:0000074; P: regulation of cell cycle; TAS.
 InterPro: IPR000719; Prot_kinase.
 InterPro: IPR002290; Ser_thr_kinase.
 Pfam: PF00069; pkinase; 1.
 ProDom: PD000001; Prot_kinase; 1.
 SMART: SM00220; S_TKc; 1.
 PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding.
 FT DOMAIN 5 288 PROTEIN KINASE.
 FT NP_BIND 11 19 ATP (BY SIMILARITY).
 FT BINDING 34 34 ATP (BY SIMILARITY).
 FT ACT_SITE 127 127 BY SIMILARITY.
 SQ SEQUENCE 358 AA: 41834 MW: 88344321F24B77C6 CRC64;

Q82bp9 yersinia pe
 P95136 mycobacteri
 Q05943 streptomyce
 P06509 spinacia ol
 P41096 lordeum val
 P17788 zea mays (m
 P17351 oryza sativ
 P1163 pisum sativ
 P11534 triticum ae
 P56791 arabidopsis
 P30065 epifagus vi
 Q9blh9 lotus japon

Query Match 24.28; Score 8; DB 1; Length 358;
 Best Local Similarity 100.0%; Pred. No. 1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 YPALGLLK 28
 DQ 259 YPALGLLK 266
 I:|||||

RESULT 2

YC YPJC_BACSU
 ID YPJC_BACSU STANDARD; PRT; 215 AA.
 AC P42978:

DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein ypjC.
 GN YPJC OR J04C.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]

SEQUENCE FROM N.A.
 RC STRAIN=168 / Marburg;
 RX MEDLINE=96349105; PubMed=8760912;
 RA Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,
 RA Serror P.;
 RT "Sequence analysis of the Bacillus subtilis chromosome region between
 RT the serA and kds loci cloned in a yeast artificial chromosome";
 RL Microbiology 142:2005-2016(1995).
 RN [2]

SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Ailoni G.,
 RA Azevedo V., Bertero M.G., Bossieres P., Bolotin A., Borchert S.,
 RA Borris R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.I.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizi A., Galleron N.,
 RA Gham S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
 RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Iwaya M., Jones L.,
 RA Joris B., Katamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.W., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogawa A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Serror S.J., Serror P., Shin B.S., Soldo B.,
 RA Sorokin A., Taccioni E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka I., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzengger T.,
 RA Winters P., Wipat A., Yamamoto K., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis";
 RL Nature 390:249-256(1997).
 RC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: TO B.SUBTILIS YIT AND YOFU.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL; AJ250712; CAB60616.1; -

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL; L38424; AAA92872.1; -
 DR EMBL; L47709; AAB38440.1; -
 DR EMBL; 299135; CAB14167.1; -
 DR PIR; C69937; C69937;
 DR S.BtList; BG11209; ypjC.
 DR InterPro; IPR003740; DUF161.
 DR Pfam; PF02588; DUF161; 1.
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 31 51 POTENTIAL.
 FT TRANSMEM 71 91 POTENTIAL.
 FT TRANSMEM 104 124 POTENTIAL.
 SQ SEQUENCE 215 AA; 23582 MW; D314CF7225F8A983 CRC64;

Query Match 21.2%; Score 7; DB 1; Length 215;
 Best Local Similarity 100.0%; Pred. No. 7;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 3 VALSLEL 9
 DQ 19 VALSLEL 16
 I:||||

RESULT 3

CDH_HUMAN
 ID CDH_HUMAN STANDARD; PRT; 224 AA.
 AC P56750;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 LT 15-SEP-2003 (Rel. 42, Last annotation update)
 CE Claudin-17.
 GN CLDN17.
 CS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

SEQUENCE FROM N.A.
 RA Keep T.J., Inglehearn C.F.;
 RA Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]

SEQUENCE FROM N.A.
 RP MEDLINE=20289799; PubMed=10830953;
 RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
 RA Park H.-S., Toyoda A., Ishii K., Toki Y., Choi D.-K., Groner Y., K.,
 RA Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,
 RA Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
 RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
 RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
 RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.B.,
 RA Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
 RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
 RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
 RA Weinmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
 RA Leinhardt R., Reinhardt R., Yaspo M.-L.;
 RA "The DNA sequence of human chromosome 21";
 RL Nature 405:311-319(2000).
 RC -!- FUNCTION: COMPONENT OF TIGHT JUNCTION (TJ) STRANDS.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: Belongs to the claudin family.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL; AJ250712; CAB60616.1; -

DR EXBL: AP001707; BAA55566.1; ..
 DR Genew: HGNC:2038; CLDN17.
 DR InterPro: IPR006187; Claudin.
 DR InterPro: IPR006184; Claudin sup.
 DR Pfam: PF00822; EMP22_Claudin.
 DR PRINTS: PR01077; CLAUDIN.
 DR PROSITE: PS01346; CLAUDIN. 1.
 KW Tight junction; Transmembrane.
 FT TRANSMEM 8 28 POTENTIAL.
 FT TRANSMEM 82 102 POTENTIAL.
 FT TRANSMEM 125 145 POTENTIAL.
 FT TRANSMEM 165 185 POTENTIAL.
 SQ SEQUENCE 224 AA; 24603 MW; 1833ED3178B7F63A CRC64;

Query Match 21.2%; Score 7; DB 1; Length 224;
 Best Local Similarity 100.0%; Pred. No. 7.2;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VAVALS 7
 DB 86 VAVALS 92

RESULT 4
 CYB_RHOVI STANDARD; PRT: 419 AA.
 ID CYB_RHOVI
 AC P81378;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DF Cytochrome b.
 GN PEB3 OR PBCB.
 OS Rhodopsinodermas viridis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Hyphomicrobiales; Blastochloris.
 OX NCBI_TaxID=1079;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SRAIN=DSK133;
 RA MEDLINE=90158506; PubMed=2560136;
 RX Verbaist J., Lang F., Gabellini N., Oesterheld D.,
 RT "Cloning and sequencing of the fcb, B and C genes encoding the
 cytochrome b/c1 complex from Rhodospseudomonas viridis";
 RL Mol. Gen. Genet. 219:445-452(1989)
 CC -i- FUNCTION: Component of the ubiquinol:cytochrome c reductase
 complex (complex III or cytochrome b-c1 complex), which is a
 respiratory chain that generates an electrochemical potential
 coupled to ATP synthesis (By similarity).
 CC -i- COFACTOR: Binds two heme groups non-covalently. Heme 1 (or B1 or
 b562) is low-potential and absorbs at about 552, and heme 2 (or B2
 or b566) is high-potential and absorbs at about 566 (By
 similarity).
 CC -i- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b.
 cytochrome c1 and the Rieske protein (By similarity).
 CC -i- SIMILARITY: Belongs to the cytochrome b family.
 PR: JQ0346; JQ0346.
 DR InterPro: IPR005798; Cytb_b6_C.
 DR InterPro: IPR005797; Cytb_b6_N.
 DR Pfam: PF00032; Cytochrome_b_c1.
 DR Pfam: PF00033; Cytochrome_b_N_1.
 DR PROSITE: PS00192; CYTOCHROME_B_HEME; ..
 DR PROSITE: PS00193; CYTOCHROME_B_OO; 1.
 KW Electron transport; Respiratory chain; Heme; Transmembrane.
 FT METAL 96 96 IRON 1 (HEME B562 AXIAL LIGAND).
 FT METAL 110 110 IRON 2 (HEME B566 AXIAL LIGAND).
 FT METAL 197 197 IRON 1 (HEME B562 AXIAL LIGAND).
 FT METAL 211 211 IRON 2 (HEME B566 AXIAL LIGAND).
 SQ SEQUENCE 419 AA; 47237 MW; 202C0C3D25B2A6DB CRC64;

Query Match 21.2%; Score 7; DB 1; Length 419;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 FLWLGA 14
 DB 351 FLWLGA 357

RESULT 5
 CG51_HUMAN STANDARD; PRT: 469 AA.
 ID CG51_HUMAN
 AC Q9Y512; Q9Y512; Q9Y512; Q9Y512; Q9Y512; Q9Y512;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Protein CGI-51.
 GN CG151.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANT ILF-345.
 RC MEDLINE=20272150; PubMed=10810093;
 RA Lai C.-H., Chou C.-Y., Ch'ang L.-Y., Liu C.-S., Lin W.-C.,
 RT "Identification of novel human genes evolutionarily conserved in
 Caenorhabditis elegans by comparative proteomics";
 RL Genome Res. 10:703-713(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Nagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Yananabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saifu K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuho Y.,
 RA Ninomiya K., Iwayanagi T.;
 RC "NEDO human cDNA sequencing project";
 RN Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20057165; PubMed=10591208;
 RA Dunham I., Hunt A.R., Collins J.E., Bruskiewicz R., Beare D.M.,
 RA Ciamp M., Shink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
 RA Baguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
 RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
 RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
 RA Clegg S.W., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
 RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
 RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
 RA Evans K.L., Fey J.M., Fleming K.K., French L., Garner A.A.,
 RA Gilbert J.G.R., Goward M.E., Graham D.V., Griffiths M.N.D., Hall C.,
 RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
 RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
 RA Laird G.K., Langford C.F., Levensha M.A., Lloyd C., Lloyd D.M.,
 RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., McCann O.T.,
 RA McElay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.C.,
 RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
 RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
 RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
 RA Soderlund C., Spraggon L., Steward C.A., Sulston J.E., Swann R.M.,
 RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
 RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
 RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
 RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
 RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
 RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
 RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.L.,
 RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
 RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
 RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
 RA Zhan M., Zhang G., Chissoe S., Murray J., Miller N., Minx P.,
 RA Zhan M., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
 RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,

RA Hinds K., Kemp X., Latreille P., Layman D., Ozersky P., Rohlfing T.,
RA Scheet P., Walker C., Wamsley A., Wohldmann P., Pepin K., Nelson J.,
RA Korf I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R.,
RA Emanuel B.S., Shaikh T., Kurahashi H., Saitta S., Budarf M.L.,
RA McDermid H.E., Johnson A., Wong A.C., Morrow B.E., Edelmann L.,
RA Kim U.J., Shizuya H., Simon M., Dumanski C.P., Peyrard M., Kedia D.,
RA Seroussi E., Fransson I., Tapia I., Bruder C.E., O'Brien K.P.,
RA Wilkinson P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,
RA Tiliahun Y., Wright H.;
RA "The DNA sequence of human chromosome 22.";
RT Nature 402:489-495(1999).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph. Muscle, and Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Schmeier C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore J., Max S.I., Wang C., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gararatne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Nelson E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko V., Houfard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Greenwood J., Schmitt J., Myers R.M.,
RA Butterfield A., Schein J.E., Jones S.J.N., Marra M.A.,
RA Schnerch A., Schen J.E., Jones S.J.N., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:15899-15903(2002).
CC -!- SIMILARITY: BELONGS TO THE UPF0140 (CGI-51) FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AK001087; BAA91498.1;
DR EMBL: AF151809; RAD34046.1;
DR EMBL: AL035398; CAB51401.1;
DR EMBL: BC007830; AAH07830.1;
DR EMBL: BC011680; AAH11681.1;
DR EMBL: BC015200; AAH15200.1;
DR InterPro: IPR000184; Bac_surfAg_D15.
DR Pfam: PF01103; Bac_surface_Ag; 1.
KW Polymorphism.
FT VARIANT 345 345 V -> I (IN dbSNP:8418).
FT /FTIQ-VAR-013768.
FT FT CONFLICT 110 110 D -> G (IN REF. 2).
FT FT CONFLICT 368 372 WAGGL -> IGRRW (IN REF. 1).
FT FT CONFLICT 371 371 MISSING (IN REF. 4; AAH07830).
SQ SEQUENCE 469 AA; 51962 MW; 4F687D27A12092EF CRC64;
Query Match 21.2%; Score 7; DB 1; Length 469;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 14 ADRFYLG 20
DB 334 ADRFYLG 340
|||||||
RESULT 6
CEAB_ECOLI STANDARD; PRT; 510 AA.
ID CEAB_ECOLI

AC PC5819;
DC 01-NOV-1988 (Rel. 09, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Colicin B.
GN CBA.
OS Escherichia coli.
OG Plasmid ColBM-pPfl66
CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC Enterobacteriaceae; Escherichia.
OX NCB:TaxID=562;
KN [1]
KP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=87250309; PubMed=2439491;
KA Schramm E., Mende J., Braun V., Kamp R.M.;
RT "Nucleotide sequence of the colicin B activity gene cba: consensus
pentapeptide among tonB-dependent colicins and receptors.";
RL J. Bacteriol. 169:3350-3357(1987).
CC -!- FUNCTION: THIS COLICIN IS A CHANNEL-FORMING COLICIN. THIS CLASS OF
TRANSMEMBRANE TOXINS DEPOLARIZE THE CYTOPLASMIC MEMBRANE, LEADING
TO DISSIPATION OF CELLULAR ENERGY.
CC -!- FUNCTION: COLICINS ARE POLYPEPTIDE TOXINS PRODUCED BY AND ACTIVE
AGAINST ESCHERICHIA COLI AND CLOSELY RELATED BACTERIA.
CC -!- MISCELLANEOUS: THIS COLICIN REQUIRES TONB FOR ITS UPTAKE.
CC -!- SIMILARITY: HIGH IN THE N-TERMINAL FIRST 300 AMINO ACID RESIDUES
WITH COLICIN D.
CC -!- SIMILARITY: WITH OTHER CHANNEL FORMING COLICINS, HIGHEST WITH
COLICIN A.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M16816; AAA98063.1;
DR PIR: A27089; IKECBB.
DR HSP: P04480; ICOL.
DR InterPro: IPR000293; Channel_colicin.
DR InterPro: IPR003058; Cloacin.
DR InterPro: IPR000531; TonB_boxC.
DR Pfam: PF03515; Cloacin; 1.
DR Pfam: PF01024; Colicin; 1.
DR PRINTS: PR00280; CHANNELCOLICIN.
DR PRINTS: PR01295; CLOACIN.
DR ProDom: PD002657; Channel_colicin; 1.
DR PROSITE: PS00276; CHANNEL_COLICIN; 1.
DR PROSITE: PS00430; TONB_DEPENDENT_REC_1; 1.
DR PROSITE: PS01556; TONB_DEPENDENT_REC_2; FALSE_NEG.
KW Antibiotic; Bacteriocin; Plasmid; Transmembrane; TonB box.
FT INIT_MET 0 0
FT SITE 16 23 TONB BOX.
FT TRANSMEM 454 474 POTENTIAL.
FT TRANSMEM 476 496 POTENTIAL.
SQ SEQUENCE 510 AA; 54732 MW; 6E4B972CF19245F1 CRC64;
Query Match 21.2%; Score 7; DB 1; Length 510;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 AVALSIF 8
DB 457 AVALSIF 463
|||||||
RESULT 7
ITR1_SCHPO STANDARD; PRT; 575 AA.
ID ITR1_SCHPO
AC Q10286;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update);
 DE Myo-inositol transporter 1.
 GN ITR1 OR SPAC4F8.15 OR SPAC7D4.G1.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=968 F50;
 RX MEDLINE=98228265; PubMed=9560432;
 RA Niederberger C., Graub R., Schweingruber A.-M., Fankhauser H.,
 RA Rusu M., Politeia M., Edenharter L., Schweingruber M.F.,
 RA "Exogenous inositol and genes responsible for inositol transport are
 RT required for mating and sporulation in Schizosaccharomyces pombe.";
 RL Curr. Genet. 33:255-261(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Scouras J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gollins S., Goble A., Hamlin A., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean C.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeli C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkhardt G., Aert R., Rebhen J., Glynnoprez B.,
 RA Weltjens I., Vanneste E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzyl K., Langer I., Beck A., Leinach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Reyson A., Thoden G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., de Ray F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong C., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Bartell B.G., Nurse P.,
 RA "The genome sequence of Schizosaccharomyces pombe."
 RJ Nature 415:871-880(2002).
 CC -!- FUNCTION: TRANSPORTER FOR MYO-INOSITOL.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X98622; CA67211.1;
 DR EMBL: Z94530; CAB11061.1;
 DR EMBL: Z99532; CAB16718.1;
 DR PIR: T43400; T43400.
 DR GeneDB SPombe: SPAC4F8.15;
 DR InterPro: IPR007114; MIPS.
 DR InterPro: IPR005828; Sub_transporter.
 DR InterPro: IPR005829; Sug_transporter.
 DR InterPro: IPR003663; Sugar_transp.
 DR Pfam: PF00083; sugar tr. 1.
 DR PRINTS: PR00171; SUGTRANSPO.
 DR TIGRFAMs: TIGR00679; SP; 1.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 2.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; FALSE_NEG.

KW Transmembrane; Sugar transport; Glycoprotein.
 FT DOMAIN 1 86 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 87 107 1 (POTENTIAL).
 FT DOMAIN 108 129 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 130 150 2 (POTENTIAL).
 FT DOMAIN 151 155 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 157 177 3 (POTENTIAL).
 FT DOMAIN 178 186 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 187 207 4 (POTENTIAL).
 FT DOMAIN 208 215 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 216 236 5 (POTENTIAL).
 FT DOMAIN 237 245 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 247 267 6 (POTENTIAL).
 FT DOMAIN 268 349 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 350 376 7 (POTENTIAL).
 FT DOMAIN 371 376 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 377 397 8 (POTENTIAL).
 FT DOMAIN 398 400 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 401 421 9 (POTENTIAL).
 FT DOMAIN 422 441 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 442 462 10 (POTENTIAL).
 FT DOMAIN 463 486 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 487 507 11 (POTENTIAL).
 FT DOMAIN 508 510 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 511 531 12 (POTENTIAL).
 FT DOMAIN 532 575 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 432 432 N-LINKED (GLCNAC...)(POTENTIAL).
 SQ SEQUENCE 575 AA; 62757 MW; 3B7C5EFF86C596AE CRC64;
 Query Match 21.2%; Score 7; DB 1; Length 575;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 8 FLGWLGA 14
 DE 519 FLGWLGA 525
 RESULT 8
 LINK2_RAI
 ID LINK2_RAT STANDARD: PRT: 638 AA.
 AC P53670;
 DT 01-OCT-1996 (Rel. 34, Created)
 DI 01-OCT-1996 (Rel. 34, Last sequence update)
 DI 28-FEB-2003 (Rel. 41, Last annotation update)
 DE LIM domain kinase 2 (tc 2.7.1.-) (LINK2-2).
 GN LINK2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS LINK2A; LINK2B; LINK2C AND LINK2D).
 RC STRAIN=Wistar; TISSUE=Brain;
 RX MEDLINE=95380177; PubMed=7651734;
 RA Nunoue K., Ohashi K., Okano I., Mizuno K.;
 RA "LINK-1 and LINK-2, two members of a LIM motif-containing protein
 RT kinase family.";
 RI Oncogene 11:701-710(1995).
 CC -!- FUNCTION: DISPLAYS SPRINE/THREONINE-SPECIFIC PHOSPHORYLATION OF
 CC MYELIN BASIC PROTEIN AND HISTONE (MBP) IN VITRO (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=4;
 CC Name=LINK2A;
 CC IsoID=P53670-1; Sequence=Displayed;
 CC Name=LINK2B;
 CC IsoID=P53670-2; Sequence=VSP_003128;
 CC Name=LINK2C;
 CC IsoID=P53670-3; Sequence=VSP_003128, VSP_003129, VSP_003130;
 CC Name=LINK2D;
 CC IsoID=P53670-4; Sequence=VSP_003128, VSP_003131;
 CC -!- TISSUE SPECIFICITY: FOUND IN VARIOUS TISSUES AT MODERATE LEVELS.

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: AE000291; AAC75061.1; ALT_INIT.
 DR EMBL: D90838; BAA15825.1; ALT_INIT.
 DR EMBL: D90839; BAA15832.1; ALT_INIT.
 DR EMBL: U24429; AAB47869.1; -.
 DR HSSP: P07305; LSND.
 DR EcoGene: EG12686; flu.
 DR InterPro: IPR006315; Autotransport.
 DR InterPro: IPR005546; Autotransporter.
 DR InterPro: IPR004899; Pertactin.
 DR Pfam: PF03797; Autotransporter; 1.
 DR Pfam: PF03212; Pertactin; 1.
 DR TIGRFAMs: TIGR01414; autotrans_bar1; 1.
 KW Outer membrane; Signal; Complete proteome.
 FT SIGNAL 1 52
 FT CHAIN 53 551 ANTIGEN 43 ALPHA CHAIN.
 FT CHAIN 552 1039 ANTIGEN 43 BETA CHAIN.
 FT VARIANT 2 2 K -> N (IN STRAIN ML 308-225).
 FT VARIANT 41 42 SL -> FF (IN STRAIN ML 308-225).
 FT VARIANT 46 46 T -> K (IN STRAIN ML 308-225).
 FT VARIANT 157 157 W -> L (IN STRAIN ML 308-225).
 FT VARIANT 188 188 V -> F (IN STRAIN ML 308-225).
 FT VARIANT 303 305 AIN -> STI (IN STRAIN ML 308-225).
 FT VARIANT 320 320 A -> T (IN STRAIN ML 308-225).
 FT VARIANT 372 372 N -> Q (IN STRAIN ML 308-225).
 FT VARIANT 493 493 E -> V (IN STRAIN ML 308-225).
 FT VARIANT 497 497 S -> N (IN STRAIN ML 308-225).
 FT VARIANT 585 585 H -> Y (IN STRAIN ML 308-225).
 FT VARIANT 709 709 E -> K (IN STRAIN ML 308-225).
 FT VARIANT 721 721 M -> T (IN STRAIN ML 308-225).
 FT VARIANT 751 753 GHL -> SHF (IN STRAIN ML 308-225).
 FT VARIANT 803 603 S -> P (IN STRAIN ML 308-225).
 FT VARIANT 815 615 A -> V (IN STRAIN ML 308-225).
 FT VARIANT 824 824 C -> S (IN STRAIN ML 308-225).
 FT VARIANT 829 835 LNLVHTS -> MNLVYNA (IN STRAIN ML 308-225).
 FT VARIANT 845 847 QGT -> LGA (IN STRAIN ML 308-225).
 FT VARIANT 855 855 S -> T (IN STRAIN ML 308-225).
 FT VARIANT 888 888 Q -> L (IN STRAIN ML 308-225).
 FT VARIANT 1025 1025 S -> I (IN STRAIN ML 308-225).
 FT CONFLICT 61 63 EIV -> IIT (IN REP. 5).
 SQ SEQUENCE 1039 AA; 10584; MW: 51706647.08; EPR60 CRC64;

Query Match 21.2%; Score 7; DB 1; Length 1039;
 Best Local Similarity 100.0%; Pred No. 22;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VAVALSJ 7
 Db 36 VAVALSJ 42

RESULT 10
 PSAG_PEA STANDARD; PRT; 39 AA.
 AC P20120;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Photosystem I reaction center subunit V (PSI-G) (Photosystem I 9 kDa protein) (Fragment).
 DE PSAG.
 GN Pisum sativum (Garden pea).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
 OX NCBI_TaxID=3888;
 RN [1]
 RP SEQUENCE.

RX MEDLINE=88137587; PubMed=3277857;
 RA Dunn P.J., Packman L.C., Pappin D., Gray J.C.;
 FT "N-terminal amino acid sequence analysis of the subunits of pea
 RT photosystem I.";
 RL FEBS Lett. 228:157-161(1988).
 CC -!- FUNCTION: NOT YET KNOWN.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast
 CC thylakoid membrane (Probable).
 CC -!- SIMILARITY: Belongs to the psag / psak family.
 DE PIR; S00318; S00318.
 DE InterPro: IPR000549; PSI_Psag/K.
 DE Pfam: PF01241; PSI_Psag; 1.
 DE PROSITE: PS01026; PHOTOSYSTEM_I_Psag; PARTIAL.
 KW Photosynthesis; Photosystem I; Transmembrane; Chloroplast; Thylakoid.
 FT NON_TER 39
 SQ SEQUENCE 39 AA; 4397 MW; 00E8A9E219D91A3 CRC64;

Query Match 18.2%; Score 6; DB 1; Length 39;
 Best Local Similarity 100.0%; Pred No. 19;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LSLFLG 10
 Db 14 LSLFLG 19

RESULT 11
 TY13_HUMAN STANDARD; PRT; 58 AA.
 ID TY13_HUMAN
 AC Q9BZ97;
 DI 28-FEB-2003 (Rel. 41, Created)
 DI 28-FEB-2003 (Rel. 41, Last sequence update)
 DI 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Transcript Y13 protein.
 GN TY13 OR TY13.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Kuroda-Kawaguchi T., Skaletsky H., Minx P.J., Brown L.G., Rozen S.,
 RA Wilson R.K., Waterston R.H., Paig D.C.;
 RC "The DNA sequence of the human Y chromosome."
 RC Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: AF332242; AAK13492.1; -.
 CC Genew: HGNC:18494; TTY13.
 DE Transmembrane.
 KW TRANSMEM 17 37 POTENTIAL.
 SQ SEQUENCE 58 AA; 6256 MW; F714A679F062DFE CRC64;

Query Match 18.2%; Score 6; DB 1; Length 58;
 Best Local Similarity 100.0%; Pred No. 26;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LSLFLG 10
 Db 24 LSLFLG 29

RESULT 12
 YG12_BACHD STANDARD; PRT; 111 AA.
 ID YG12_BACHD

Q9KCF9;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein BH1612.
 GN BH1612.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A. / JCM 9153;
 RC STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fuji F., Hirata C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis."
 RL Nucleic Acids Res. 28:4317-4331(2000).
 CC -!- SIMILARITY: BELONGS TO THE UPF0131 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AP001512; BAB05331.1;
 DR PIR: D83851; D83851.
 DR InterPro: IPR005347; JPF0131.
 DR Pfam: PF03674; UPF0131.1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 111 AA; 12782 MW; 24730C75CC2096FE CRC64;

 Query Match 18.28; Score 6; DB 1; Length 111;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 19 LGYPAL 24
 Db 22 LGYPAL 27

 RESULT 13
 ID PSCA_HUMAN STANDARD; PRT; 123 AA.
 AC O43653;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Prostate stem cell antigen precursor.
 GN PSCA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Prostatic carcinoma;
 RX MEDLINE=98132661; PubMed=9465086;
 RA Reiter R.F., Gu Z., Watabe T., Thomas G., Szigeti K., Davis E.,
 RA Wahl M., Nistari S., Yamashiro J., le Beau M.M., Loda M., Witte O.N.;
 RT "Prostate stem cell antigen: a cell surface marker overexpressed in
 RT prostate cancer."
 RL Proc. Natl. Acad. Sci. U.S.A. 95:1735-1740(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Urothelial;
 RA Bahrenberg G., Joost H.G.;
 RT "Serial analysis of the gene expression of a highly differentiated
 RT urothelial tumor.";

Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 [3]
 RN TISSUE SPECIFICITY.
 RX MEDLINE=20180504; PubMed=10713670;
 RA Gu Z., Thomas G., Yamashiro J., Shintaku I.P., Dorey F., Raitano A.,
 RA Witte O.N., Said J.W., Loda M., Reiter R.F.;
 RT "Prostate stem cell antigen (PSCA) expression increases with high
 RT gleason score, advanced stage and bone metastasis in prostate
 RT cancer."
 RL Oncogene 19:1288-1296(2000).
 CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PROSTATE (BASAL, SECRETORY
 CC AND NEUROENDOCRINE EPITHELIUM CELLS). ALSO FOUND IN BLADDER
 CC (TRANSITIONAL EPITHELIUM), PLACENTA (TROPHOBLASTS), STOMACH
 CC (NEUROENDOCRINE CELLS), COLON (NEUROENDOCRINE CELLS) AND KIDNEY
 CC (COLLECTING DUCTS); OVEREXPRESSED IN PROSTATE CANCERS AND
 CC EXPRESSION IS CORRELATED WITH TUMOR STAGE, GRADE AND ANDROGEN-
 CC INDEPENDENCE. HIGHLY EXPRESSED IN PROSTATE CANCER BONE METASTASES.
 CC -!- SIMILARITY: Contains 1 UPAR/Ly6 domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF043498; AAC39657.1;
 DR EMBL: AJ297436; CAB57347.1;
 DR GenBank; HGNC:9500; PSCA.
 DR MIM: 602470;
 DR GO: GO:0005886; C:plasma membrane; TAS.
 DR GO: GO:0008222; F:tumor antigen; TAS.
 DR InterPro: IPR003632; Ly6_CD59.
 DR InterPro: IPR001526; Ly6_UPAR.
 DR Pfam: PFC00021; UPAR_Ly6; 1.
 DR ProDom: PD003128; Ly6_CD59; 1.
 DR SMART: SM00134; LU; 1.
 DR PROSITE: PS00983; Ly6_UPAR; Membrane; GPI-anchor.
 KW Signal; Antigen; Glycoprotein; Membrane; GPI-anchor.
 FT SIGNAL 1 20 POTENTIAL
 FT CHAIN 21 95 PROSTATE STEM CELL ANTI-GEN.
 FT PROPEP 96 123 REMOVED IN MATURE FORM (POTENTIAL).
 FT DOMAIN 21 95 UPAR/LY6.
 FT DISULFID 23 48 BY SIMILARITY.
 FT DISULFID 26 35 BY SIMILARITY.
 FT DISULFID 41 66 BY SIMILARITY.
 FT DISULFID 70 86 BY SIMILARITY.
 FT DISULFID 87 92 BY SIMILARITY.
 FT CARBOHYD 40 40 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 83 83 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 93 93 N-LINKED (GLCNAC...) (POTENTIAL).
 FT LIPID 95 95 GPI-ANCHOR (POTENTIAL).
 SQ SEQUENCE 123 AA; 12912 MW; 3FC1271742D657FA CRC64;

 Query Match 18.28; Score 6; DB 1; Length 123;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 22 PALGLL 27
 Db 111 PALGLL 116

 RESULT 14
 ID GTH2_MORSA STANDARD; PRT; 139 AA.
 AC Q91121;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Gonadotropin beta-II chain precursor (GTH-II-beta).
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Urothelial;
 RA Bahrenberg G., Joost H.G.;
 RT "Serial analysis of the gene expression of a highly differentiated
 RT urothelial tumor.";

OS Morone saxatilis (Striped bass).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percidae;
 OC Moronidae; Morone.
 OX NCBI_TaxID=34816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pituitary;
 RX MEDLINE=96020549; PubMed=8546811;
 RA Hassin S., Elizur A., Zohar Y.;
 RI "Molecular cloning and sequence analysis of striped bass (Morone
 saxatilis) gonadotropin-I and -II subunits.";
 RL J. Mol. Endocrinol. 15:23-35(1995).
 CC -!- FUNCTION: INVOLVED IN GAMETOGENESIS AND STEROIDOGENESIS.
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta chain.
 CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L35096; AAC38019.1; ALT_INIT.
 CC HSSP: P01233; 1XUL.
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR002400; GF_CysKnot.
 DR InterPro: IPR001545; Gly_hormoneB.
 DR Pfam: PF00007; Cys_knot; 1.
 DR PRINTS: PR03438; GFCYSKNOT.
 DR SMART: SM00068; GHb; 1.
 DR PROSITE: PS00261; GLYCO_HORMONE_BETA_1; 1.
 DR PROSITE: PS00689; GLYCO_HORMONE_BETA_2; 1.
 KW Hormone; Glycoprotein; Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 139 GONADOTROPIN BETA-1 CHAIN.
 FT DISULFID 30 78 BY SIMILARITY.
 FT DISULFID 44 93 BY SIMILARITY.
 FT DISULFID 47 131 BY SIMILARITY.
 FT DISULFID 55 109 BY SIMILARITY.
 FT DISULFID 59 111 BY SIMILARITY.
 FT DISULFID 114 121 BY SIMILARITY.
 FT CARBOHYD 34 34 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 139 AA; 15555 MW; 1A88B742F81AF76 CRC64;
 Query Match: 18.2%; Score 6; DB 1; Length 139;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 LSLFLG 10
 Db 11111
 6 LSLFLG 11
 RESULT 15
 PSAG_ARATH
 ID PSAG_ARATH STANDARD; PRI: 160 AA.
 AC Q95N7; Q42310;
 DT 28-FEB-2003 (Rel. 41; Created)
 DT 28-FEB-2003 (Rel. 41; Last sequence update)
 DT 28-FEB-2003 (Rel. 41; Last annotation update)
 DE Photosystem I reaction center subunit V, chloroplast precursor (PS:-
 G).
 DE PSAG OR ATIG55670 OR F20N2.3 OR F20N2.33 OR F20N2.21.
 GN Arabidopsis thaliana (Mouse-ear cress).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 Query Match: 18.2%; Score 6; DB 1; Length 160;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 LSLFLG 10
 Db 11111
 6 LSLFLG 11

[1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=cv. Columbia;
 RC MEDLINE=22016719; PubMed=1130712;
 RX White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
 KA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Egtu P., Feildlyum T.V., Feng J.-B., Fong B., Fujii C.Y.,
 RA Giller J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
 PA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 KA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzilli A.,
 RA Mietscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RC "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 thaliana.";
 RL Nature 408:816-820(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 FT "Riken Arabidopsis full length cDNA clones (RAFLs) sequenced by the
 RT SSP consortium (Salk/Stanford/PGECC).";
 RN Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 27-127 FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Bardet C., Dabos P., Tremousaygue D., Lescure B.;
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: NOT YET KNOWN.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast
 CC thylakoid membrane (Probable).
 CC -!- SIMILARITY: Belongs to the psag / psak family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AJ245630; CAB52748.1;
 CC EMBL: AC002328; AAC83069.1;
 CC EMBL: AF326870; AAC41452.1;
 CC EMBL: AF324710; AAC40061.1;
 CC EMBL: AF339692; AAC00374.1;
 CC EMBL: Z37244; CAAB5530.1;
 CC In-Pr: IPRC00549; PSI_Psag/K.
 DR Pfam: PF01241; PSI_Psag; 1.
 DR PROSITE: PS01026; PHOTOSYSTEM_I_Psag; 1.
 KW Photosynthesis; Photosystem I; Transmembrane; Chloroplast; Thylakoid;
 KW Transit peptide.
 FT TRANSIT 1 60 CHLOROPLAST (BY SIMILARITY).
 FT CHAIN 61 160 PHOTOSYSTEM I REACTION CENTER SUBUNIT V.
 FT TRANSMEM 66 86 POTENTIAL.
 FT TRANSMEM 129 149 POTENTIAL.
 SQ SEQUENCE 160 AA; 17085 MW; 6EF23E57C7C50760 CRC64;
 Query Match: 18.2%; Score 6; DB 1; Length 160;

Best Local Similarity 100.0%; Pred. No. 54;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 LSLFLG 10
jllll
Db 74 LSLFLG 79

Search completed: September 26, 2003, 17:37:20
Job time : 4.93192 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2003, 17:29:45 ; Search time 17.1319 seconds
(without alignment)
497.069 Million cell updates/sec

Title: US-09-852-100A-2_COPY_185_217
Perfect score: 33
Sequence: 1 VAAVLSFLGWLGWGADRFYLGYPALGLIKFCVVG 33

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protist:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	100.0	207	4 Q9BX74	Q9BX74 Homo sapiens
2	33	100.0	208	11 Q99MB3	Q99MB3 Mus musculus
3	8	24.2	80	4 Q9H046	Q9H046 Homo sapiens
4	8	24.2	178	5 Q9W2H1	Q9W2H1 Drosophila
5	8	24.2	195	5 Q95QZ5	Q95QZ5 Caenorhabditis
6	8	24.2	221	4 Q9H651	Q9H651 Homo sapiens
7	8	24.2	230	11 Q9D156	Q9D156 Mus musculus
8	8	24.2	247	4 Q9BRN9	Q9BRN9 Homo sapiens
9	8	24.2	261	11 Q8BJ83	Q8BJ83 Mus musculus
10	8	24.2	284	5 Q9H4H5	Q9H4H5 Drosophila
11	7	21.2	38	11 Q8CAM1	Q8CAM1 Mus musculus
12	7	21.2	150	16 Q8EV17	Q8EV17 Mycoplasma
13	7	21.2	159	10 Q9LXV9	Q9LXV9 Arabidopsis
14	7	21.2	180	16 Q8FPV7	Q8FPV7 Corynebacterium
15	7	21.2	194	16 Q92QW4	Q92QW4 Rhizobium
16	7	21.2	219	16 Q9RX10	Q9RX10 Deinococcus

17	7	21.2	221	12 Q9ITN3	Q9ITN3 Tupia harrisi
18	7	21.2	251	11 Q8BS94	Q8BS94 Mus musculus
19	7	21.2	259	17 Q8PVR4	Q8PVR4 Methanosaeta
20	7	21.2	269	16 Q8RRR3	Q8RRR3 Thermotoga
21	7	21.2	273	16 Q83115	Q83115 Treponema
22	7	21.2	314	2 Q8RLA5	Q8RLA5 Escherichia
23	7	21.2	317	17 Q57894	Q57894 Pyrococcus
24	7	21.2	329	12 Q8U2J1	Q8U2J1 Cercopithecus
25	7	21.2	349	16 Q9PLE9	Q9PLE9 Chlamydia
26	7	21.2	358	16 Q8CUN9	Q8CUN9 Oceanobaculum
27	7	21.2	401	2 Q9S458	Q9S458 Salmonella
28	7	21.2	407	16 Q8ED47	Q8ED47 Shewanella
29	7	21.2	415	16 Q915C8	Q915C8 Pseudomonas
30	7	21.2	444	2 Q9RQ60	Q9RQ60 Rhizobium
31	7	21.2	448	2 Q24842	Q24842 Acinetobacter
32	7	21.2	451	11 Q9QU14	Q9QU14 Mus musculus
33	7	21.2	466	2 Q87185	Q87185 Streptococcus
34	7	21.2	466	2 Q9ALW7	Q9ALW7 Streptococcus
35	7	21.2	466	2 Q9AFH1	Q9AFH1 Streptococcus
36	7	21.2	466	2 Q93113	Q93113 Streptococcus
37	7	21.2	466	2 Q9AQ17	Q9AQ17 Streptococcus
38	7	21.2	466	16 Q8E500	Q8E500 Streptococcus
39	7	21.2	466	16 Q8D2E2	Q8D2E2 Streptococcus
40	7	21.2	469	11 Q8BGH2	Q8BGH2 Mus musculus
41	7	21.2	492	16 Q9RUA0	Q9RUA0 Deinococcus
42	7	21.2	510	17 Q59101	Q59101 Pyrococcus
43	7	21.2	547	16 Q9R2J4	Q9R2J4 Deinococcus
44	7	21.2	572	10 Q8VYV8	Q8VYV8 Arabidopsis
45	7	21.2	606	13 Q8QHMO	Q8QHMO Xenopus

ALIGNMENTS

RESULT 1

Q9BX74 ID Q9BX74 PRELIMINARY: PRT: 207 AA.
AC Q9BX74;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DI 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Beta-amyloid binding protein.
GN BAP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA Kajkowski E.M., Lo C.F., Ning X., Walker S., Sofia H.J., Wang W.,
Edris W., Chanda P., Wagner E., Vile S., Ryan K., McHendry-Rinde H.,
Smith S.C., Wood A., Rhodes E.J., Kennedy J.D., Bard J.,
Jacobson J.S., Ozenberger B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein.
Containing a G Protein Activation Module.";
RE J. Biol. Chem. 275:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Testis;
RA Strausberg R.;
RI Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR ENBL; AF353990; AAK35064.1; -;
CR ENBL; BC029486; AAH29486.1; -;
KW Signal.
FT SIGNAL 1 37 POTENTIAL.
SQ SEQUENCE 207 AA; 22326 MW; A5590FD7AECDF292 CRC64;

Query Match 100.0%; Score 33; DB 4; Length 207;
Best Local Similarity 100.0%; Pred. No. 2.1e-24;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VAAVLSFLGWLGWGADRFYLGYPALGLIKFCVVG 33

Db 123 VAVALSFLWLGADRFYLYGYPALGLLKFTVIG 155
|||||

RESULT 2

Q99MB3 PRELIMINARY; PRT; 208 AA.
AC Q99MB3;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Beta-amyloid binding protein.
GN BBP.
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=21276355; PubMed=11278849;
RA Kajkowski E.M., Lo C.F., Ning X., Walker S., Sofia H.J., Wang W.,
RA Edris W., Chandra P., Wagner E., Vile S., Ryan K., McHenry-Rinde R.,
RA Smith S.C., Wood A., Rhodes K.J., Kennedy J.D., Bard J.,
RA Jacobsen J.S., Ozenberger B.A.,
RT "Beta-Amyloid Peptide-Induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RL J. Biol. Chem. 276:18748-18756(2001).
DR EMBL; AF353993; AA33067.1;
DR MGD; MG12137022; Bbp.
SQ SEQUENCE 208 AA; 22271 MW; 92A7932163F4F04C CRC64;

Query Match 100.0%; Score 33; DB 11; Length 268;
Best Local Similarity 100.0%; Pred. No. 2.2e-24;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VAVALSFLWLGADRFYLYGYPALGLLKFTVIG 33

|||||
Db 124 VAVALSFLWLGADRFYLYGYPALGLLKFTVIG 156
|||||

RESULT 3

Q9H046 PRELIMINARY; PRT; 80 AA.
AC Q9H046;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Hypothetical protein: (fragment).
GN DKFZP667C1011.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph node;
RA Koehler K., Beyer A., Mewes H.W., Weil R., Wiemann S.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL512689; CAC21647.1;
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 80 AA; 8699 MW; 8B6BE768235C58D CRC64;

Query Match 24.2%; Score 8; DB 4; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 GADRFYLYG 20

|||||
Db 33 GADRFYLYG 40
|||||

RESULT 4
Q9W2H1 PRELIMINARY; PRT; 178 AA.
AC Q9W2H1;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE CG10795 protein (LD27358P).
GN CG10795.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Alt J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bereman B.P., Bhardani D., Bolshakov S.,
RA Borokova J., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Fertaz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.D., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclob J.M.,
RA Patuzzolo M., Pittman G.S., Pan S., Poillard J., Puri V., Reese H.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun F.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brockstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuncio J., Paclob J., Parag V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003453; AAF46720.1;
DR EMBL; AY061343; AAL28891.1;
DR Flybase; FBgn0034626; CG10795.
SQ SEQUENCE 178 AA; 19896 MW; 17C41166607ACC03 CRC64;

Query Match 24.2%; Score 8; DB 5; Length 178;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Caps 0;

```

QY      15 DREYLYGYP 22
      11111111
DB      108 DREYLYGYP 115

RESULT 5
Q95Q25 PRELIMINARY: PRT: 195 AA.
AC Q95Q25;
DT 01-DEC-2001 (TrEMBLrel. 19, Created);
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update);
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update);
DE Hypothetical 21.2 kDa protein.
GN C41D11.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peleoderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=95069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Gattung S., Magoi L.;
RT "The sequence of C. elegans cosmid C41D11.";
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003740; AAL08031.1;
DR WormPep; C41D11.9; CE29489.
KW Hypothetical protein.
SQ SEQUENCE 195 AA: 21203 MW: 35945E407F134DAE CRC64;

Query Match 24.2%; Score 8; DB 5; Length 195;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      13 GADRYLYG 20
      11111111
DB      148 GADRYLYG 155

RESULT 6
Q9H651 PRELIMINARY: PRT: 221 AA.
AC Q9H651;
DT 01-MAR-2001 (TrEMBLrel. 16, Created);
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update);
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update);
DE Hypothetical protein FLJ22604 (BBP-like protein 2).
GN BLP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,
RA Nakamura Y., Isono T., Sugano S.;
RT "NEO human cDNA sequencing project.";

```

```

RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP Kajkowski E.M., Lo C.F., Ning X., Walker S., Sofia H.J., Wang W.,
RA Edris W., Charada P., Wagner E., Ville S., Ryan K., McHenry-Rinde B.,
RA Smith S.C., Wood A., Rhodes K.J., Kennedy J.D., Bard J.,
RA Jacobsen J.S., Ozenberger B.A.;
RC "Beta-amyloid peptide-induced apoptosis regulated by a novel protein
RC containing a G protein activation module.";
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RC Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK026257; BAB15415.1;
DR EMBL; AF353992; AAK35066.1;
DR EMBL; BC008873; AAH08873.1;
KW Hypothetical protein.
SQ SEQUENCE 221 AA: 24410 MW: 92151D6EF6363D74 CRC64;

Query Match 24.2%; Score 8; DB 4; Length 221;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      13 GADRYLYG 20
      11111111
DB      174 GADRYLYG 181

RESULT 7
Q9D156 PRELIMINARY: PRT: 230 AA.
AC Q9D156;
DT 01-JUN-2001 (TrEMBLrel. 17, Created);
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update);
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update);
DE 1110025109rik protein (RIKEN cDNA 1110025109 gene).
GN 1110025109RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RC MEDLINE=210855660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Frieschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schram L.M., Staubil F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido I., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Jyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Morgone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilmink L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kotsuki S.,
RA Hayashizaki Y.;
RC "Functional annotation of a full-length mouse cDNA collection.";
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RA Strausberg R.;
RC Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK003917; BAB23075.1;

```

```

DR EMBL: BC024620; AAH24620.1;
DR MGI: 1915884; 1110025109Rik.
SQ SEQUENCE 230 AA; 25639 MW; 396D650D8BE99A5 CRC64;

Query Match      24.2%; Score 8; DB 11; Length 230;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;

QY      13 GADRFYLG 20
DB      183 GADRFYLG 190
      |||||
      |||||

RESULT 8
Q9BRN9 ID Q9BRN9 PRELIMINARY; PRT: 247 AA.
AC Q9BRN9
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DSC-2001 (TrEMBLrel. 19, Last annotation update)
DE Similar to hypothetical protein FIJ22604.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DDRJ databases.
DR EMBL: BC006150; AAH06150.1;
KW Hypothetical protein.
SQ SEQUENCE 247 AA; 27161 MW; CE1D0D9C53CDF73C CRC64;

Query Match      24.2%; Score 8; DB 4; Length 247;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;

QY      13 GADRFYLG 20
DB      200 GADRFYLG 207
      |||||
      |||||

RESULT 9
Q8BJ83 ID Q8BJ83 PRELIMINARY; PRT: 261 AA.
AC Q8BJ83
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Similar to BBP-like protein 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Forelimb;
RX MEDLINE=22354583; PubMed=12466851;
RA The FANTOM Consortium.
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs."
DR EMBL: AK077858; BAC37037.1;
SQ SEQUENCE 261 AA; 28880 MW; 70346780D3CF5CDB CRC64;

Query Match      24.2%; Score 8; DB 11; Length 261;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;

QY      13 GADRFYLG 20
      |||||

```

```

DB      214 GADRFYLG 221

RESULT 10
Q9U4H5 ID Q9U4H5 PRELIMINARY; PRT: 284 AA.
AC Q9U4H5; Q9W361;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE BCDA.GH02974 (ALMONDEX) (AMX protein).
GN AMX OR BCDA:GH02974 OR CG12127.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,
RA Agbayani A., Arcaina T.T., Baxter E., Blazej R.G., Butenhoff C.,
RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
RA Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,
RA Houston K.A., Hummasti S.R., Kim E., Li P., Moshrefi M., Pacleb J.M.,
RA Park S., Sequeira A., Sethi H., Snir E., Svirska R.R., Weinburg T.,
RA Celisner S.E.;
RT "Full Length Drosophila melanogaster cDNA sequence."
RC Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Ovary;
RA Micheliou M.A.E., Remillieux N.C., Randsholt N.B.;
RT "Characterization of almondex."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.K., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Balieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandal D., Bolshakov S.,
RA Borokova D., Botchan M.K., Bouck J., Brokstein P., Brottier P.,
RA Surtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Durkin B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Morkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

```


RA Gibos R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 DR EMBL: AF181623; AAD53409.1; -.
 DR EMBL: AF211797; AAF36924.2; -.
 DR EMBL: AE003446; AAF46474.2; -.
 DR FlyBase: FBgn0000077; amx.
 DR InterPro: IPR001304; Lectin_C.
 DR PROSITE: PS00615; C_TYPE_LECTIN_1; 1.
 SQ SEQUENCE 284 AA; 31364 MW; 8FB8FF573AC851 CRC64;

Query Match 24.2%; Score 8; DB 5; Length 284;
 Best Local Similarity 100.0%; Pred. No. 6.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 GADRFYLG 20
 DB 237 GADRFYLG 244
 |||||

RESULT 11
 Q8CAM1 PRELIMINARY; PRT; 38 AA.
 AC Q8CAM1;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE LIM motif-containing protein kinase 2 (fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Hypothalamus;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium.
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 DR EMBL: AK038530; BAC30030.1; -.
 FT NON_TER 1
 SQ SEQUENCE 38 AA; 4188 MW; 2DB363A49415D42 CRC64;

Query Match 21.2%; Score 7; DB 11; Length 38;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ALSIFLG 10
 DB 3 ALSIFLG 9
 |||||

RESULT 12
 Q8EVI7 PRELIMINARY; PRT; 150 AA.
 AC Q8EVI7;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN MYPE5770.
 OS Mycoplasma penetrans.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=28227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HF-2;
 RX MEDLINE=22354719; PubMed=12466555;
 RA Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K.,
 RA Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;
 RT "The complete genomic sequence of *Mycoplasma penetrans*, an

RT Intracellular bacterial pathogen in humans."
 RL Nucleic Acids Res. 30:5293-5300(2002).
 DR EMBL: AP004172; BAC44367.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 150 AA; 17289 MW; 0FF392EB9F32F1F1 CRC64;

Query Match 21.2%; Score 7; DB 16; Length 150;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 GADRFYL 19
 DB 70 GADRFYL 76
 |||||

RESULT 13
 Q5LXV9 PRELIMINARY; PRT; 159 AA.
 AC Q5LXV9;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN T15B3.160 OR AT3G44020/T15B3.160.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Jordan N., Bangert S., Wiedelmann R., Voss H., Unseld M., Mewes H.W.,
 RA Rudd S., Lemcke K., Mayer K.F.X., Quettier F., Salanoubat M.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,
 RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
 RA Hayashizaki Y., Shinozaki K.;
 RT "Arabidopsis thaliana full-length cDNA."
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL163975; CAB88130.1; -.
 DR EMBL: AK118643; BAC43239.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 159 AA; 17059 MW; A3F0A7B03B5AR78A CRC64;

Query Match 21.2%; Score 7; DB 10; Length 159;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ALSIFLG 10
 DB 94 ALSIFLG 100
 |||||

RESULT 14
 Q8FPV7 PRELIMINARY; PRT; 180 AA.
 AC Q8FPV7;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Conserved hypothetical protein.
 GN CE1361.
 OS Corynebacterium efficiens.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=152794;

```

RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RA  Kawarabayashi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
RA  Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,
RA  Usuda Y., Sugimoto S.;
RT  "The entire genomic sequence of Corynebacterium efficiens YS-314."
RL  Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR  EMBL: AP005218; BAC18191.1;
KW  Hypothetical protein; Complete proteome.
SQ  SEQUENCE 180 AA; 19455 MW; EA6BF8CA54FCD761 CRC64;

```

```

Query Match          21.2%; Score 7; DB 16; Length 180;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy  22 PALGLLK 28
    |||||
Db  133 PALGLLK 139

```

RESULT 15

```

O920W4
ID  Q920W4      PRELIMINARY:      PRT: 194 AA.
AC  Q920W4;
DT  01-DEC-2001 (TrEMBLrel. 19, Created)
DT  01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT  01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE  Hypothetical protein acvB.
GN  ACV8 OR F01182 OR SMC00613.
OS  Rhizobium meliloti (Sinorhizobium meliloti).
OC  Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC  Rhizobiaceae; Sinorhizobium.
OX  NCBI_TaxID=382;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=1021;
RX  MEDLINE=21396507; PubMed=11481430;
RA  Capella D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Barut C.,
RA  Boistard P., Becker A., Routry M., Cadieu F., Drearo S., Gloux S.,
RA  Godrie T., Coffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA  Pohl T., Portetelle D., Pehier A., Purnelle B., Ramsperger U.,
RA  Renard C., Thebaud P., Vandenbol M., Weidner S., Gallibert F.;
RT  "Analysis of the chromosome sequence of the legume symbiont
RL  Sinorhizobium meliloti strain 1021."
RL  Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001);
DR  EMBL: AL591786; CAC45761.1;
DR  InterPro; IPR000734; Lipase
DR  InterPro; IPR000379; Ser_gstrs_site.
DR  PROSITE; PS00120; LIPASE_SER; 1.
KW  Hypothetical protein; Complete proteome.
SQ  SEQUENCE 194 AA; 21357 MW; AABE97B2813F5E42 CRC64;

```

```

Query Match          21.2%; Score 7; DB 16; Length 194;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy  9 LGWLGAD 15
    |||||
Db  113 LGWLGAD 119

```

```

Search completed: September 26, 2003, 17:41:38
Job time : 20.1319 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

CM protein - protein search, using sw model

Ran on: September 26, 2003, 17:28:58 : Search time 102.289 seconds
(without alignments)
313.452 Million cell updates/sec

Title: US-09-852-100A-2_COPY_68_269

Perfect score: 202

Sequence: 1 PSGPSAPEAVTARLVGLWLP.....TRDRLSLINETFRKQLYV 204

Scoring table: OLIGO
Gapop 60.0 , Gapext 50.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000600

Post-processing: Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

- 1: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
- 24: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	202	100.0	269	20 AAW94291	Human beta-amyloid
2	202	100.0	269	21 AAY70759	Human beta-amyloid
3	202	100.0	269	24 AAE33877	Human BHP-1 prote.
4	133	65.8	139	20 AAY12358	Human 5' EST seque
5	111	55.0	162	20 AAY36021	Extended human 5'
6	82	40.6	148	20 AAY12426	Human 5' EST seque
7	43	21.3	100	23 AAY97631	RNA polymerase II
8	17	8.4	68	24 AAE33878	Human BHP-1 protei
9	9	4.5	225	22 ABB11574	Human secreted pro

10	8	4.0	24	23	ABU77688	New peptide vector
11	8	4.0	27	23	ABB81176	Signal sequence ba
12	8	4.0	27	23	ABG78990	Cell penetrating p
13	8	4.0	27	23	AAE23685	Fluorescently labe
14	8	4.0	27	23	ABU77687	New peptide vector
15	8	4.0	27	23	AAU78348	Signal sequence ba
16	8	4.0	27	24	ABG75506	Signal sequence ba
17	8	4.0	27	24	ABB82542	Signal sequence ba
18	8	4.0	178	22	ABB65236	Drosophila melanog
19	8	4.0	221	21	RAY70761	Human beta-amyloid
20	8	4.0	221	24	ABR47818	Human secreted pro
21	8	4.0	221	24	ABR48493	Human Amyloid Apop
22	8	4.0	221	24	ABR00112	Human gene 102 enc
23	8	4.0	222	20	AAU27658	Human secreted pro
24	8	4.0	230	21	AAU08645	DNA encoding a mur
25	8	4.0	247	21	AAU08641	A human TANGO 223
26	8	4.0	247	21	AAU08656	A human TANGO 223
27	8	4.0	247	21	AAU08657	A human TANGO 223
28	8	4.0	247	21	AAU08658	A human TANGO 223
29	8	4.0	284	22	ABB59014	Drosophila melanog
30	8	4.0	358	23	ABP62954	Human polypeptide
31	8	4.0	1638	23	ABP62827	Drosophila melanog
32	7	3.5	15	23	AAU97632	RNA polymerase II
33	7	3.5	18	23	AAU70969	M. tuberculosis Rv
34	7	3.5	24	23	AAU75279	Leucine-rich repea
35	7	3.5	35	22	AAU08325	Human immune/haema
36	7	3.5	55	23	ABP08508	Human ORFX protein
37	7	3.5	57	22	ABG16850	Novel human diagn
38	7	3.5	58	22	AAU63743	Propionibacterium
39	7	3.5	63	23	ABP04761	Human ORFX protein
40	7	3.5	76	22	AAU51626	Propionibacterium
41	7	3.5	88	23	AAU14446	Protein of human s
42	7	3.5	95	22	AAU44112	Propionibacterium
43	7	3.5	102	23	AAU70953	M. tuberculosis Rv
44	7	3.5	106	21	ABG01646	Human secreted pro
45	7	3.5	108	22	AAU60957	Propionibacterium

ALIGNMENTS

RESULT 1
AAW94291
ID AAW94291 standard; Protein: 269 AA.
AC AAW94291;
DE 27-APR-1999 (first entry)
EE Human beta-amyloid peptide-binding protein (BBP).
FF Beta-amyloid peptide binding protein; BBP; beta-amyloid protein; BAP;
GW human; Alzheimer's disease.
XX Homo sapiens.
XX Key Location/Qualifiers
FH Region 68..269
FT /note: "specifically claimed fragment having
FT beta-amyloid peptide binding activity."
XX
FN W09846636-A2.
XX
PP 22-OCT-1998.
PP 14-APR-1998; 98WO-US07462.
XX
PR 16-APR-1997; 97US-0064583.
XX
PA (AMHP) AMERICAN HOME PROD CORP.
XX
PI Hard JA. Jacobsen JS, Kajkowski EM, Ozenberger BA;
PI Walker SG;

```
XX WPI: 1999-080736/07.
DR N-PSDB: AAX05735.
XX
PT Polynucleotide encoding beta-amyloid peptide binding protein - used
PT to identify inhibitors of beta-amyloid peptide for treating
PT Alzheimer's disease
XX
PS Claim 7: Pages 43-44; 59pp; English.
XX
CC The present sequence represents a beta-amyloid peptide binding protein
CC (BBP). The polynucleotide comprising the entire BBP nucleotide sequence
CC of clone BBP-1 is deposited under the accession number ATCC 98617. The
CC polynucleotide comprising a fragment of BBP (nucleotides 202-807 of the
CC full length BBP) of clone pK196 is deposited as ATCC 98399. Host cells
CC transformed with a vector comprising the BBP nucleic acid are used in a
CC the recombinant production of the protein. The protein can be used in a
CC method for diagnosing a disease characterised by aberrant expression of
CC human beta-amyloid protein (BAP). The protein can also be used in a
CC method for screening for compounds which regulate expression of a BAP
CC binding protein. The proteins, antibodies and identified compounds can be
CC used in the treatment or prevention of Alzheimer's disease.
XX
SQ Sequence 269 AA:
Query Match 100.0%; Score 202; DB 20; Length 269;
Best Local Similarity 100.0%; Pred. No. 7e-193;
Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PSGPSAPEAVTARLVGVLFVSVITGPGAVATSGAGEESLKCEDLKVGQYICKDKPKIND 60
Db 68 PSGPSAPEAVTARLVGVLFVSVITGPGAVATSGAGEESLKCEDLKVGQYICKDKPKIND 127
QY 61 ATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFKPISCRNVNGSYKVAV 120
Db 128 ATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFKPISCRNVNGSYKVAV 187
QY 121 ALSLFGLWLGADRFLYGLYPALGLKFCITGFCGIGSLIDFILISMQIVGPSGSSYIDY 180
Db 188 ALSLFGLWLGADRFLYGLYPALGLKFCITGFCGIGSLIDFILISMQIVGPSGSSYIDY 247
QY 181 YGTRLRSLTINETFRKTLQYP 202
Db 248 YGTRLRSLTINETFRKTLQYP 269
RESULT 2
AAY70759
ID AAY70759 standard; Protein: 269 AA.
AC AAY70759;
XX
XX 24-JUL-2000 (first entry)
XX
XX Human beta-amyloid peptide (BAP) binding protein, BBP1.
XX
KW Beta-amyloid peptide binding protein; BBP; BAP; tumour; suppressor;
KW G-protein coupled receptor; GPCR; integral membrane protein; antigen;
KW neuronal cell; nonhuman primate; NHP; G-protein signalling pathway;
KW apoptosis; immunogen; therapeutic; treatment; prevention; diagnostic.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Domain 177..198
FT Domain /label= Transmembrane_domain_1
FT Domain 199..201
FT /label= DRF_motif
FT /note= Substitution of the Arg abrogates protection"
FT Domain 213..238
FT /label= Transmembrane_domain_2
XX
XX WO200022125-A2.
PN
```

```
XX 20-APR-2000.
XX
XX 13-OCT-1999; 99WO-US21621.
XX
XX 13-OCT-1998; 98US-0104104.
XX
XX (AMHP ) AMERICAN HOME PROD CORP.
XX
XX Ozenberger BA, Kajkowski EM, Lo CF;
XX WPI: 2000-3:7982/27.
XX N-PSDB: AA252369.
XX
XX Novel G-protein-coupled receptor-like proteins and polynucleotides
XX useful for regulating apoptosis, comprises integral membrane protein
XX traversing the membrane twice -
XX
XX Example 1: Page 62-63; 68pp; English.
XX
CC The present sequence is the beta-amyloid peptide (BAP) binding protein-1
CC (BBP1). It is an integral membrane protein, that traverse the membrane
CC twice. It is related to G protein-coupled receptor (GPCR) protein
CC superfamily. It interacts with G-alpha proteins and regulates the
CC activity of G-protein signalling pathways. Bap genes are widely expressed
CC in neuronal cells of nonhuman primate (NHP) brain and overexpressed in
CC some tumours. It functions as a suppressor of apoptosis induction. BAP
CC proteins are used as immunogens to raise antibodies, useful as
CC therapeutics and as antigens in solid phase assays. They are also useful
CC as reagents to identify molecules which effect the interaction of BBP and
CC a cloned protein, that are useful in the treatment or prevention of
CC diseases associated with apoptosis. The polynucleotides are useful for
CC diagnostics.
XX
XX Note: In claim 5, the patent claims an amino acid sequence from figure 2.
XX However, figure 2 does not contain any sequence. It is inferred from the
XX disclosure that the figure 2 sequence refers to BBP1 protein, shown in
XX this sequence.
XX
XX Sequence 269 AA:
Query Match 100.0%; Score 202; DB 21; Length 269;
Best Local Similarity 100.0%; Pred. No. 7e-193;
Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PSGPSAPEAVTARLVGVLFVSVITGPGAVATSGAGEESLKCEDLKVGQYICKDKPKIND 60
Db 68 PSGPSAPEAVTARLVGVLFVSVITGPGAVATSGAGEESLKCEDLKVGQYICKDKPKIND 127
QY 61 ATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFKPISCRNVNGSYKVAV 120
Db 128 ATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFKPISCRNVNGSYKVAV 187
QY 121 ALSLFGLWLGADRFLYGLYPALGLKFCITGFCGIGSLIDFILISMQIVGPSGSSYIDY 180
Db 188 ALSLFGLWLGADRFLYGLYPALGLKFCITGFCGIGSLIDFILISMQIVGPSGSSYIDY 247
QY 181 YGTRLRSLTINETFRKTLQYP 202
Db 248 YGTRLRSLTINETFRKTLQYP 269
RESULT 3
AAE33877
ID AAE33877 standard; Protein: 269 AA.
XX
XX AAE33877;
XX
XX 02-MAY-2003 (first entry)
XX
XX Human BBP-1 protein.
XX
XX Human; beta-amyloid peptide-binding protein; BAP; Abeta; betaAP; BBP;
XX Alzheimer's disease; AD; transgenic; transgenic animal; gene therapy;
KW
```

KW neuroprotective; nootropic.

XX Homo sapiens.

XX W0200290499-A2.

XX 14-NOV-2002.

XX 06-MAY-2002; 2002WO-US14223.

XX 09-MAY-2001; 2001US-0852100.

XX (AMHP) WYETH.

XX Ozenberger BA, Bard JA, Kajkowski EM, Jacobson JS, Walker SG;

PI Sofia RJ, Howland DS;

XX WPI; 2003-120537/11.

XX N-PSDB; AAD51940.

XX New human beta-amyloid peptide-binding protein, useful for diagnosing
PT and/or treating diseases associated with aberrant expression of
PI beta-amyloid peptide, e.g. Alzheimer's disease .

XX Claim 4; Page 84-85; 85pp; English.

XX The present invention relates to novel human beta-amyloid peptide (BAP;
CC Abeta, betaAP)-binding (BBP) proteins and polynucleotides encoding such
CC proteins. BBP sequences are useful to diagnose and/or treat diseases
CC associated with aberrant expression of human BAP such as Alzheimer's
CC disease (AD). They are used to generate transgenic animals. Sequences
CC of the invention are also used in gene therapy. The present sequence
CC is human BBP-1 protein.

XX Sequence 269 AA;

Query Match 100.0%; Score 202; DB 24; Length 269;

Best Local Similarity 100.0%; Pred. No. 7e-193;

Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSGSAPAVTARLVGVLFVSVTTGPGAVATSAGGESLKCEDLKVGQYICKDKPKIND 63

DB 68 PSGSAPAVTARLVGVLFVSVTTGPGAVATSAGGESLKCEDLKVGQYICKDKPKIND 127

QY 61 ATOEPVNCNTNTAHVSCFPAPNITCKOSSGNETHFTGNEVGFFKPISCRNVNGYSYKVA 120

DB 128 ATOEPVNCNTNTAHVSCFPAPNITCKOSSGNETHFTGNEVGFFKPISCRNVNGYSYKVA 187

QY 121 ALSFLGLWLGADRFYLGYPALGLLKFCTVGFCSGLIDFIDLSMQIVGPSDSSYIIDY 180

DB 188 ALSFLGLWLGADRFYLGYPALGLLKFCTVGFCSGLIDFIDLSMQIVGPSDSSYIIDY 247

QY 181 YGTRLTRLSITNETRKTQIYP 202

DB 248 YGTRLTRLSITNETRKTQIYP 259

RESULT 4

AAV12358

ID AAY12358 standard; Protein; 139 AA.

XX AAY12358;

XX 17-JUN-1999 (first entry)

XX Human 5' EST secreted protein SEQ ID NO:399.

XX Human; secreted protein; EST; expressed sequence tag; diagnosis;

XX forensic; gene therapy; chromosome mapping; signal peptide;

XX upstream regulatory sequence; cytokine activity; cell proliferation;

XX differentiation; haematopoiesis regulation; tissue growth regulation;

XX reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;

XX thrombolytic; anti-inflammatory; tumour inhibition.

XX Homo sapiens.

XX W09906548-A2.

XX 11-FEB-1999.

XX 31-JUL-1998; 98WO-1801222.

XX 01-AUG-1997; 97US-0905135.

XX (GEST) GENSET.

XX Duclert A, Dumas Milne Edwards J, Lacroix B;

XX WPI; 1999-153776/13.

XX N-PSDB; AAX41191.

XX New nucleic acids encoding human secreted proteins obtained from
PT cDNA libraries prepared from e.g. liver, ovary, brain, prostate,
PI kidney, lung, umbilical cord, placenta and colon tissue

XX Claim 27; Page 714-715; 824pp; English.

XX AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for
CC human secreted proteins, and encode the proteins given in AAY12261 to
CC AAY12514, respectively. The proteins given represent the signal peptide
CC and an N-terminal fragment of a secreted protein. The nucleic acid
CC sequences can be used for producing secreted human gene products. They
CC can also be used to develop products for diagnosis and therapy. The
CC proteins obtained may have cytokine activity, cell
CC proliferation/differentiation activity, haematopoiesis regulating
CC activity, tissue growth regulating activity, reproductive hormone
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptide can be used for
CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell.

XX Sequence' 139 AA;

Query Match 65.88; Score 133; DB 20; Length 139;

Best Local Similarity 100.0%; Pred. No. 2.3e-124;

Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SGPSAPAVTARLVGVLFVSVTTGPGAVATSAGGESLKCEDLKVGQYICKDKPKINDA 61

DB 7 SGPSAPAVTARLVGVLFVSVTTGPGAVATSAGGESLKCEDLKVGQYICKDKPKINDA 66

QY 62 TOEPVNCNTNTAHVSCFPAPNITCKOSSGNETHFTGNEVGFFKPISCRNVNGYSYKVA 120

DB 67 TOEPVNCNTNTAHVSCFPAPNITCKOSSGNETHFTGNEVGFFKPISCRNVNGYSYKVA 126

QY 122 LSFLGLWLGADRF 134

DB 127 LSFLGLWLGADRF 139

RESULT 5

AAV126021

ID AAY36021 standard; Protein; 162 AA.

XX AAY36021;

XX 13-SEP-1999 (first entry)

XX Extended human secreted protein sequence, SEQ ID NO. 406.

XX Secreted protein; human; cytokine; cellular proliferation; cell movement;

XX cellular differentiation; immune system regulator; anti-inflammatory;


```

XX PD 16-JAN-2002.
XX PF
XX PS 30-JUN-2000; 2000CN-0116963.
XX PR 30-JUN-2000; 2000CN-0116963.
XX PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX PI Mao Y, Xie Y;
XX PS WPI: 2002-340664/38.
XX DR N-PSDB; ABK52558.
XX SQ
PT Polypeptide-RNA polymerase II subunit I; and polynucleotide for coding
PT it.
XX
XX PS Claim 1; Page 29; 32pp; Chinese.
XX PR
XX CC This invention relates to the DNA and protein sequences of a novel
XX polypeptide-RNA polymerase II subunit I protein. The invention also
XX comprises a process for preparing the polypeptide of the invention by
XX DNA recombination, the application of the polypeptide in treating
XX diseases such as cancer, human immunodeficiency virus (HIV) infection,
XX etc, the antagonist of the polypeptide and its medical action, and the
XX application of the said polynucleotide are disclosed. The present
XX sequence represents the RNA polymerase II subunit I protein of the
XX invention.
XX SQ Sequence 100 AA:
Query Match 21.3%; Score 43; DB 23; Length 100;
Best Local Similarity 100.0%; Pred. No. 8.6c-35;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 9 AVTARLVGLVWFVSVTGPGCAVATSGAGGESLKCEDLKVGQY 51
DB 14 AVTARLVGLVWFVSVTGPGCAVATSGAGGESLKCEDLKVGQY 56
RESULT 8
AAE33878
ID AAE33878 standard; Protein; 68 AA.
XX AC AAE33878;
XX DT 02-MAY-2003 (first entry)
XX DE Human BBP-1 protein fragment.
XX KW Human: beta-amyloid peptide-binding protein; BAP; Abeta; betaAP; BAP;
XX Alzheimer's disease; AD; transgenic; transgenic animal; gene therapy;
XX neuroprotective; nontropic.
XX OS Homo sapiens.
XX PN WO200290499-A2.
XX PR 14-NOV-2002.
XX PD
XX PF 06-MAY-2002; 2002WO-US14223.
XX PR 09-MAY-2001; 2001US-0852100.
XX PS (AMHP) WYETH.
XX PI Ozenberger BA, Bard JA, Kajkowski EM, Jacobsen JS, Walker SS;
XX Sofia HJ, Howland DS;
XX WPI: 2003-120537/11.
XX DR N-PSDB; AAD51978.
XX PT New human: beta-amyloid peptide-binding protein, useful for diagnosing

```

```

PT and/or treating diseases associated with aberrant expression of
PT beta-amyloid peptide, e.g. Alzheimer's disease.
XX
XX PS Example 11; Fig 9; 85pp; English.
XX PR
XX CC The present invention relates to novel human beta-amyloid peptide (BAP;
XX Abeta, betaAP)-binding (BBP) proteins and polynucleotides encoding such
XX proteins. BBP sequences are useful to diagnose and/or treat diseases
XX associated with aberrant expression of human BAP such as Alzheimer's
XX disease (AD). They are used to generate transgenic animals. Sequences
XX of the invention are also used in gene therapy. The present sequence
XX is human BBP-1 protein fragment.
XX SQ Sequence 68 AA:
Query Match 6.4%; Score 17; DB 24; Length 68;
Best Local Similarity 100.0%; Pred. No. 4.9e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 150 GFCGIGSLIDFILISMQ 166
DB 1 GFCGIGSLIDFILISMQ 17
RESULT 9
ABH11574
ID ABH11574 standard; peptide; 225 AA.
XX AC ABH11574;
XX DT 11-JAN-2002 (first entry)
XX DE Human secreted protein homologue, SEQ ID NO:1944.
XX KW Human: cytokine; cell proliferation; cell differentiation; growth factor;
XX haematopoiesis regulation; tissue growth; immunomodulator; activin;
XX inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
XX proliferation; metastasis; cancer; tumour; haematopoietic disorder;
XX myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
XX chronic inflammatory condition; proliferative retinopathy;
XX atherosclerosis; coronary heart disease; arterial ischaemia;
XX bone disorder; osteoporosis; vascular growth disorder;
XX tissue regeneration; wound healing; infection; immune disorder;
XX cell culture; drug screening; gene therapy; antiinflammatory;
XX antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
XX cytoslatic; osteopathic; vasotropic; cardiac; virucide; antibacterial;
XX antifungal; vulnery; antiulcer.
XX OS Homo sapiens.
XX PN WO200157188-A2.
XX PR 09-AUG-2001.
XX PD
XX PF 05-FEB-2001; 2001WO-US03800.
XX PR 03-FEB-2000; 2000US-0456914.
XX PR 27-APR-2000; 2000US-0569875.
XX PS (HYSE-) HYSEQ INC.
XX PI Tang Y, Liu C, Drmanac RT;
XX WPI: 2001-457740/49.
XX DR N-PSDB; ABA08818.
XX PR
XX PT Human proteins and DNA encoding sequences useful for preventing,
XX treating or ameliorating a medical condition in a mammalian subject
XX e.g. arthritis and cancer.
XX PS Claim 20; Page 214; 1963pp; English.
XX SQ Sequences ABH10981-ABH12330 represent 1350 novel human polypeptides, and

```

sequences ABA08225-ABA09574 represent nucleic acids encoding them. The invention also relates to vectors and recombinant host cells comprising a nucleotide of the invention, methods of producing the novel polypeptides, antibodies against the polypeptides, methods of detecting the nucleotides or polypeptides in a sample, and methods of identifying compounds which bind to polypeptides of the invention. Although novel, many of the polypeptides of the invention have homology to known proteins, thereby giving an insight into their probable biological activities, and hence potential therapeutic applications. The polypeptides of the invention may have various activities, including cytokine, cell proliferation or cell differentiation activities; stem cell growth factor activity; haematopoietic regulatory activity; tissue growth activity; immunomodulatory activity; activin- or inhibin-related activities; chemotactic or chemokinetic activities; haemostatic, thrombotic or thrombolytic activities; receptor or ligand activities; or may be involved in oncogenesis, cancer cell proliferation or metastasis. Depending on their biological activities, polypeptides and nucleotides of the invention are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Such conditions include cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell disorders), chronic inflammatory conditions (e.g., asthma or arthritis), proliferative retinopathy, atherosclerosis, coronary heart disease, arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal vascular growth. Polypeptides involved with tissue regeneration and repair (or nucleic acids encoding them) may be used to promote wound healing (e.g., of burns, incisions and ulcers), while those with immunomodulatory activities may be used in the treatment of viral, bacterial and fungal infections in addition to immune disorders. Polypeptides with growth factor activity may be used in cell cultures to promote cell growth. For example, such polypeptides may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease or accidental damage. The polypeptides and nucleotides may also be used in the diagnosis of the above conditions, and in drug screening techniques. The present sequence represents a novel human polypeptide of the invention.

XX Sequence 225 AA;

Query Match 4.5%; Score 9; DB 22; Length 225;

Best Local Similarity 100.0%; Pred. No. 1.4;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 LGADRFYLG 137

DB 176 LGADRFYLG 184

RESULT 10
ARB77688

ID ARB77688 standard; peptide; 24 AA.

XX AC ARB77688;

DI 01-JUL-2002 (first entry)

DE New peptide vector#4.

XX Intracellular delivery; transfection agent; cancer; infectious disease; peptide vector.

OS Synthetic.

XX WO200210201-A2.

PD 07-FEB-2002.

XX 26-JUL-2001; 2001WO-US23406.

PF 31-JUL-2000; 2000US-221932P.

PR (ACTI-) ACTIVE MOTIF.

PA (CNRS) CENT NAT RECH SCI.

XX Divida G, Morris M, Mery J, Heitz F, Fernandez J, Archdeacon J;
PI Horndorff K;

XX WPI; 2002-329441/35.

XX Transfection agent that comprises a peptide comprising hydrophobic and
PT hydrophilic domain and having amino acid residues of specified length
PI is useful for a non-covalent association with and transport of a
PT heterologous compound into a cell -

XX Example 2; Page 61; 156pp; English.

XX The invention relates to a transfection agent comprising a peptide of
CC about 16-30 amino acids in length. Peptides of the invention comprise
CC a hydrophobic domain, a hydrophilic domain, optionally a spacer sequence
CC between the domains and a functional group conjugated to at least one
CC terminal of the peptide. Peptides of the invention are useful for a
CC non-covalent association with and transport of a heterologous compound
CC into a cell. They are also useful for promoting the cellular
CC internalisation of at least one member e.g. peptide, proteins,
CC antibodies, their derivatives and/or conjugates. They may form part of a
CC pharmaceutical composition to deliver the compound selected from a
CC diagnostic or therapeutic compound, to treat at least one condition: such
CC as cancer or an infectious disease, or which targets a cancerous cell or
CC pathogen-infected cell and to deliver a peptide or inhibitor that
CC disrupts the activity of the enzyme. The agent of the invention has a
CC transfection efficiency of at least 5% for at least two of the members of
CC the group of the compounds. The agent has a good delivery efficiency for
CC a broad spectrum of compounds and cell types, has a low toxicity, are
CC easy to handle and easy to formulate in conjunction with the many
CC different compound types that it can deliver. The peptides are serum
CC sensitive, thus they code particularly well for systemic and/or localised
CC in patients. The current sequence represents a new amphipathic peptide
CC vector of the invention that contains a cationic nuclear localisation
CC sequence.

XX Sequence 24 AA;

Query Match 4.0%; Score 8; DB 23; Length 24;

Best Local Similarity 100.0%; Pred. No. 1.7;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 LFLGWLGA 131

DB 3 LFLGWLGA 10

RESULT 11
ABB81176

ID ABB81176 standard; peptide; 27 AA.

XX AC ABB81176;

XX 25-NOV-2002 (first entry)

XX Signal sequence based peptide 1 MPS peptide.

XX Red blood cell; RBC; virus; vertebrate; therapeutic; tissue imaging;
KW qp41; membrane translocation.

XX Human immunodeficiency virus type 1.

OS Rhesus macaque polyoma virus.

XX WO2002050416-A1.

PD 08-AUG-2002.

XX 01-FEB-2002; 2002WO-GB00437.

XX 01-FEB-2001; 2001GB-0002561.

PR 16-FEB-2001; 2001US-269528P.

PA (GENO-) GENDEL LTD.
 XX
 PI Mchale AP, Craig R;
 XX
 DR WPI: 2002-643355/69.
 XX
 PT Delivering agent to target site in vertebrate comprises loading red
 PT blood cell with virus or virus-like particle comprising agent,
 PT sensitizing cell, introducing cell into vertebrate and applying energy
 PT to release virus particle from cell.
 XX
 PS Disclosure: Page 55: 87pp; English.
 XX
 CC The invention relates to delivering an agent to a target site in a
 CC vertebrate. The method involves (a) loading a red blood cell (RBC) with a
 CC virus or a virus-like particle (I) comprising an agent; (b) sensitizing
 CC RBC to render it more susceptible to disruption than unsensitized RBC;
 CC (c) introducing RBC into a vertebrate, and (d) applying energy to release
 CC (I). Steps (a) and (b) may be performed in any order. (I) (RBC loaded
 CC with a virus or a virus-like particle comprising a therapeutic agent) is
 CC useful for the delivery of a therapeutic agent to a target site in a
 CC vertebrate, or in the preparation of a medicament for delivery of a
 CC therapeutic agent to a target site in a vertebrate. (II) is also useful
 CC for delivering one or more agents to a vertebrate and for treating or
 CC preventing a disease. The method is useful for delivering agents such as
 CC those useful for imaging of tissues in vivo or ex vivo, preferably for
 CC delivering an agent to a subcellular organelle such as nucleus,
 CC mitochondria, Golgi or endoplasmic reticulum. The present sequence
 CC represents a signal sequence based peptide: 1. MPS peptide, a chimera of
 CC the hydrophobic terminal domain of viral gp41 protein and the nuclear
 CC localisation signal (NLS) from the SV40 large antigen. This fragment has
 CC been found to be active in membrane translocation.
 XX
 XX Sequence 27 AA;
 SQ

Query Match 4.0%; Score 8; DB 23; Length 27;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 124 LFLGLWLGA 131
 DB 11111111
 3 LFLGLWLGA 10

RESULT 12
 ABG78990
 ID ABG78990 standard; Peptide: 27 AA.
 XX
 AC ABG78990;
 XX
 DT 15-NOV-2002 (first entry)
 XX
 DE Cell penetrating peptide Signal-peptide II.
 XX
 KW Cell penetrating peptide; cancer; tumour; melanoma; thymoma;
 KW lymphoma; sarcoma; lung cancer; non-Hodgkin's lymphoma; leukaemia;
 KW Hodgkin's lymphoma; uterine cancer; cervical cancer; bladder cancer;
 KW kidney cancer; adenocarcinoma; breast cancer; prostate cancer;
 KW ovarian cancer; pancreatic cancer; vaccine; dendritic cell;
 KW tumour infiltrating lymphocyte; TIL; human leukocyte antigen; HLA;
 KW cytostatic.
 XX
 OS Unidentified.
 XX
 PN WO200264057-A2.
 XX
 PF 22-AUG-2002.
 XX
 PD 15-FEB-2002; 2002WO-US05212.
 XX
 PF 15-FEB-2001; 2001US-265687P.
 XX
 PR (BAYU) BAYLOR COLLEGE MEDICINE.
 PA

XX Wang R;
 PI WPI: 2002-627577/67.
 XX
 DR Novel composition for treating a disease in an animal, comprises an
 PT immune effector cell and cell penetrating peptide associated with an
 PT antigen or antibody.
 XX
 FS Disclosure: Page 11: 61pp; English.
 XX
 CC The invention relates to a composition (I) comprising an immune effector
 CC cell and a cell penetrating peptide (CPP) associated with an antigen or
 CC antibody. Also included are (II) a vaccine comprising (I), CPP
 CC associated with an antigen, and a pharmaceutically acceptable carrier
 CC and (2) preparing a composition for a disease, by providing (I)
 CC and CPP associated with an antigen for disease, and introducing the
 CC antigen-associated CPP to (I), where antigen enters into the cell.
 CC The antigens are, for example, tumour antigen derived epitopes
 CC recognised by tumour infiltrating lymphocytes (TIL) of HLA (human
 CC leukocyte antigen) class I or II. The composition is useful for enhancing
 CC immunity in an animal to a disease, by administering a mature dendritic
 CC cell comprising CPP associated with an antigen to disease, to the animal,
 CC such that following the administration, animal is protected from disease,
 CC where the animal comprises both CD4+ and CD8+ T cells. It is also useful
 CC for treating a disease (e.g. cancer, tumour, melanoma, thymoma,
 CC lymphoma, sarcoma, lung cancer, non-Hodgkin's lymphoma, leukaemia,
 CC Hodgkin's lymphoma, uterine cancer, cervical cancer, bladder cancer,
 CC kidney cancer, adenocarcinoma, breast cancer, prostate cancer,
 CC ovarian cancer and pancreatic cancer). The animal is further subjected to
 CC a cancer treatment including surgery, radiation, chemotherapy or gene
 CC therapy. The administration of (I), preferably dendritic cell is prior
 CC to, subsequent to or concurrent with, the cancer treatment. The present
 CC sequence is cell penetrating peptide of the invention.
 XX
 SQ

Query Match 4.0%; Score 8; DB 23; Length 27;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 124 LFLGLWLGA 131
 DB 11111111
 3 LFLGLWLGA 10

RESULT 13
 AAE23685
 ID AAE23685 standard; peptide: 27 AA.
 XX
 AC AAE23685;
 XX
 DT 10-SEP-2002 (first entry)
 XX
 DE Fluorescently labelled RNA binding peptide #2.
 XX
 KW RNA binding protein; mRNA quantification; gene expression.
 XX
 OS Unidentified.
 XX
 PN WO200227031-A2.
 XX
 PF 04-APR-2002.
 XX
 PF 28-SEP-2001; 2001WO-US30438.
 XX
 PR 28-SEP-2000; 2000US-236407P.
 XX
 PA (CELL-) CELLOMICS INC.
 XX
 PI Busa WB;
 XX
 DR WPI: 2002-452280/48.

XX Quantifying target gene expression in living cells that possess a
 PI target gene of interest tagged with the binding site for an RNA binding
 PT protein and fluorescently labeled RNA binding polypeptide including a:
 PT RNA binding domain -
 XX
 PS Claim 45; Page 44; 51pp; English.
 XX
 CC The present invention relates to a method of quantifying the expression
 CC of target genes in living cells. The method involves providing cells that
 CC possess a target gene of interest which has been tagged with the binding
 CC site for an RNA binding protein and a fluorescently labelled RNA binding
 CC polypeptide that includes an RNA binding domain and calculating the
 CC quantity of target gene expression in the cells using fluorescence
 CC signalling techniques. The method is useful for quantifying expression
 CC of one or more target genes in living cells which comprise two or more
 CC distinct populations of cells. It is used to quantitate the expression
 CC of any target gene, including expression of protein-encoding messenger
 CC RNA genes, ribosomal RNA encoding genes and transfer RNA encoding genes
 CC so long as the RNA expression product from the target gene possesses a
 CC sequence or structure (the RNA tag) that is bound specifically by the
 CC RNA binding polypeptide being used. The present sequence is a
 CC fluorescently labelled RNA binding peptide.
 XX
 SQ Sequence 27 AA;
 QY Query Match 4.0%; Score 8; DB 23; Length 27;
 Db Best Local Similarity 100.0%; Pred. No. 1.9; Mismatches 0; Indels 0; Gaps 0;
 Matches 8; Conservative 0;
 QY 124 LFLGWLGA 131
 Db 3 LFLGWLGA 10
 I:|||||
 3 LFLGWLGA 10
 RESULT 14
 ASB77687
 ID ASB77687 standard; peptide: 27 AA.
 XX
 AC ASB77687;
 XX
 DT 01-JUL-2002 (first entry)
 XX
 DE New peptide vector#3.
 XX
 KW Intracellular delivery: transfection agent; cancer; infectious disease;
 KW peptide vector.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 7
 FT /note= "residue may be substituted with Phe"
 FT Misc-difference 23
 FT /note= "residue may be substituted with Ser"
 FT
 XX WO200210201-A2.
 PN
 XX 07-FEB-2002.
 PD
 XX 26-JUL-2001; 2001WO-US23406.
 XX
 PF 31-JUL-2000; 2000US-221932P.
 XX
 PA (ACTI-) ACTIVE MOTIF.
 PA (CNRS) CENT NAT RECH SCI.
 XX
 XX Divida G, Morris M, Mery J, Heitz F, Fernandez C, Archdeacon C;
 PI Horndorp K;
 PI
 XX WPI; 2002-329441/36.
 DR
 XX Transfection agent that comprises a peptide comprising hydrophobic and

FT hydrophilic domain and having amino acid residues of specified length
 PT is useful for a non-covalent association with and transport of a
 PT heterologous compound into a cell -
 XX
 PS Example 2; Page 61; 156pp; English.
 XX
 CC The invention relates to a transfection agent comprises a peptide of
 CC about 16-30 amino acids in length. Peptides of the invention comprise
 CC a hydrophobic domain, a hydrophilic domain, optionally a spacer sequence
 CC between the domains and a functional group conjugated to at least one
 CC terminal of the peptide. Peptides of the invention are useful for a
 CC non-covalent association with and transport of a heterologous compound
 CC into a cell. They are also useful for promoting the cellular
 CC internalisation of at least one member e.g. peptide, proteins,
 CC antibodies, their derivatives and/or conjugates. They may form part of a
 CC pharmaceutical composition to deliver the compound selected from a
 CC diagnostic or therapeutic compound, to treat at least one condition such
 CC as cancer or an infectious disease, or which targets a cancerous cell or
 CC pathogen-infected cell and to deliver a peptide or inhibitor that
 CC disrupts the activity of the enzyme. The agent of the invention has a
 CC transfection efficiency of at least 5% for at least two of the members of
 CC the group of the compounds. The agent has a good delivery efficiency for
 CC a broad spectrum of compounds and cell types, has a low toxicity, are
 CC easy to handle and easy to formulate in conjunction with the many
 CC different compound types that it can deliver. The peptides are serum
 CC sensitive, thus they bode particularly well for systemic and/or localised
 CC in patients. The current sequence represents a new amphipathic peptide
 CC vector of the invention that contains a cationic nuclear localisation
 CC sequence.
 XX
 SQ Sequence 27 AA;
 QY Query Match 4.0%; Score 8; DB 23; Length 27;
 Db Best Local Similarity 100.0%; Pred. No. 1.9; Mismatches 0; Indels 0; Gaps 0;
 Matches 8; Conservative 0;
 QY 124 LFLGWLGA 131
 Db 3 LFLGWLGA 10
 I:|||||
 3 LFLGWLGA 10
 RESULT 15
 AAU78348
 ID AAU78348 standard; Peptide: 27 AA.
 XX
 AC AAU78348;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Signal sequence based peptide I.
 XX
 FH Membrane translocation signal; signal sequence based peptide I;
 FT red blood cell vehicle; polypeptide delivery; viral gp41 protein;
 FT simian virus 40; SV40; large antigen.
 FT
 XX Rhesus macaque polyoma virus.
 CS Synthetic.
 CS
 XX WO200207752-A2.
 FN
 XX 31-JAN-2002.
 PD
 XX 24-JUL-2001; 2001WO-GH03327.
 XX
 PF 24-JUL-2000; 2000WO-GH02848.
 XX
 PR 29-AUG-2000; 2000WO-GS03056.
 PR
 PR 01-FEB-2001; 2001WO-GS0417.
 PR
 PR 16-FEB-2001; 2001US-0785802.
 XX
 XX (GENO-) GENDEL LTD.
 XX
 XX Craig K;
 PI
 XX

DR WPI: 2002-280593/32.
XX
PT Preparing a red blood cell vehicle suitable for delivering an agent to
PT a target site in a vertebrate due to loading the red blood cell with an
PT agent-membrane translocation sequence . XX
XX
PS Disclosure: Page 44; 135pp; English.
PS
XX The invention describes a method of preparing a red blood cell vehicle
CC suitable for delivering an agent to a target site in a vertebrate
CC comprising providing a red blood cell and loading the red blood cell with
CC an agent-MTS (membrane translocation sequence) conjugate. The red blood
CC cells produced may be used in the preparation of a medicament for
CC delivery of an agent to or at a target site and of one or more agents to
CC a vertebrate. The agent is actively released from the red blood cell
CC vehicle by application of a stimulus to disrupt the red blood cell
CC chimeric peptide of the hydrophobic terminal domain of the viral gp41
CC protein and the nuclear localisation signal from simian virus 40
CC (SV40) large antigen, one of the membrane translocation peptides tested
CC in the invention.
XX
SQ Sequence 27 AA:

Query Match 4.0%; Score 8; DB 23; Length 27;
Best Local Similarity 100.0%; Pred. No. 1 9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 LFLGWLGA 131
1111111
Dh 3 LFLGWLGA 10

Search completed: September 26, 2003, 17:39:24
Job time : 103.289 secs

GenCore version: 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2003, 17:36:55 : Search time 36.1021 seconds
(without alignments)
230.739 Million coll. updates/sec

Title: US-09-852-100A-2_COPY_68_269

Perfect score: 202
Sequence: 1 PGGPSAPAVTARLVGVLF.....TRTFRTSTFNETFKTKVLYP 202

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310856 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_AA:
1: /cgn2.6/ptodata/1/1aa/5A.COMB.pep:
2: /cgn2.6/ptodata/1/1aa/5B.COMB.pep:
3: /cgn2.6/ptodata/1/1aa/6A.COMB.pep:
4: /cgn2.6/ptodata/1/1aa/6B.COMB.pep:
5: /cgn2.6/ptodata/1/1aa/PCUS.COMB.pep:
6: /cgn2.6/ptodata/1/1aa/PCUSfiles.pep:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	4.0	358	US-09-411-628-11	Sequence 11, Appl
2	7	3.5	114	US-09-222-938A-13	Sequence 13, Appl
3	7	3.5	183	US-09-252-991A-21850	Sequence 21850, A
4	7	3.5	204	US-09-134-001C-5239	Sequence 5239, Ap
5	7	3.5	213	US-09-163-445B-1	Sequence 1, Appl
6	7	3.5	245	US-09-353-634-55	Sequence 55, Appl
7	7	3.5	315	US-09-393-634-56	Sequence 56, Appl
8	7	3.5	451	US-09-345-473E-45	Sequence 45, Appl
9	7	3.5	462	US-09-328-352-6888	Sequence 6888, Ap
10	7	3.5	483	US-09-252-991A-29267	Sequence 29267, A
11	7	3.5	617	US-09-345-473E-44	Sequence 44, Appl
12	7	3.5	643	US-09-252-991A-26617	Sequence 26617, A
13	7	3.5	673	US-09-063-950-2	Sequence 2, Appl
14	7	3.5	673	US-09-996-243-52	Sequence 52, Appl
15	7	3.5	707	US-09-252-991A-29401	Sequence 29401, A
16	7	3.5	2519	US-09-413-814-42	Sequence 42, Appl
17	6	3.0	23	US-08-176-508-121	Sequence 121, App
18	6	3.0	23	US-08-471-052A-121	Sequence 121, App
19	6	3.0	23	US-08-189-331-121	Sequence 121, App
20	6	3.0	23	US-08-471-939-121	Sequence 121, App
21	6	3.0	23	US-08-471-800-121	Sequence 121, App
22	6	3.0	23	US-08-471-068-121	Sequence 121, App
23	6	3.0	38	US-09-461-325-284	Sequence 284, App
24	6	3.0	46	US-09-489-847-131	Sequence 131, App
25	6	3.0	61	US-08-266-321-3	Sequence 3, Appl
26	6	3.0	61	US-08-467-527A-3	Sequence 3, Appl
27	6	3.0	61	US-08-467-528-3	Sequence 3, Appl

25	6	3.0	70	4	US-09-205-258-298	Sequence 258, App
26	6	3.0	72	3	US-08-905-223-314	Sequence 314, App
27	6	3.0	101	4	US-09-489-847-306	Sequence 306, App
28	6	3.0	117	4	US-09-228-986-111	Sequence 111, App
29	6	3.0	120	4	US-08-728-742A-1	Sequence 1, Appl
30	6	3.0	123	2	US-08-675-508-2	Sequence 2, Appl
31	6	3.0	123	3	US-09-203-939-2	Sequence 2, Appl
32	6	3.0	123	3	US-09-203-939-6	Sequence 6, Appl
33	6	3.0	123	3	US-09-251-835-2	Sequence 2, Appl
34	6	3.0	123	3	US-09-251-835-6	Sequence 6, Appl
35	6	3.0	123	3	US-09-318-503-2	Sequence 2, Appl
36	6	3.0	123	3	US-09-318-503-6	Sequence 6, Appl
37	6	3.0	123	3	US-09-318-503-6	Sequence 6, Appl
38	6	3.0	123	3	US-09-038-261A-2	Sequence 2, Appl
39	6	3.0	123	3	US-09-038-261A-6	Sequence 6, Appl
40	6	3.0	123	3	US-09-564-329A-2	Sequence 2, Appl
41	6	3.0	123	4	US-09-564-329A-6	Sequence 6, Appl
42	6	3.0	146	1	US-08-688-609-6	Sequence 6, Appl
43	6	3.0	146	3	US-09-002-832-6	Sequence 6, Appl
44	6	3.0	146	3	US-09-002-832-6	Sequence 6, Appl
45	6	3.0	146	3	US-09-002-832-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-09-411-628-11
: Sequence 11, Application US/09411628
: Patent No. 6428944
: GENERAL INFORMATION:
: APPLICANT: University of Southern California
: TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
: FILE REFERENCE: 13761-707
: CURRENT APPLICATION NUMBER: US/09/411,628
: CURRENT FILING DATE: 1999-10-01
: EARLIER APPLICATION NUMBER: US 60/102,906
: EARLIER FILING DATE: 1998-10-02
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 11
: LENGTH: 358
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-411-628-11

Query Match 4.0%, Score 8; DB 4; Length 358;
Best Local Similarity 100.0%, Pred. No. 5.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 138 YPALGLLK 145
DB 259 YPALGLLK 266

RESULT 2
US-09-222-938A-13
: Sequence 13, Application US/09222938A
: Patent No. 6437108
: GENERAL INFORMATION:
: APPLICANT: Youngman, Philip
: APPLICANT: Fritz, Chrisa
: APPLICANT: Murphy, Christopher
: APPLICANT: Guzman, Luz-Maria
: TITLE OF INVENTION: ESSENTIAL BACTERIAL GENES AND THEIR USE
: FILE REFERENCE: 07334/060001
: CURRENT APPLICATION NUMBER: US/09/222,938A
: CURRENT FILING DATE: 1998-12-30
: NUMBER OF SEQ ID NOS: 102
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 13
: LENGTH: 114
: TYPE: PRT
: ORGANISM: Streptococcus pneumoniae
US-09-222-938A-13

```
Query Match      3.5%: Score 7; DB 4; Length 114;
Best Local Similarity 100.0%: Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AVATSAG 36
Db 69 AVATSAG 75

RESULT 3
US-09-252-991A-21850
; Sequence 5239, Application US/09252091A
; Patent No. 6561795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21850
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21850

Query Match      3.5%: Score 7; DB 4; Length 183;
Best Local Similarity 100.0%: Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSGPSAP 7
Db 131 PSGPSAP 137

RESULT 4
US-09-134-001C-5239
; Sequence 5239, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/054,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5239
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5239

Query Match      3.5%: Score 7; DB 4; Length 204;
Best Local Similarity 100.0%: Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AVATSAG 36
Db 115 AVATSAG 121

RESULT 5
```

```
US-09-163-445B-1
; Sequence 1, Application: US/09163445B
; Patent No. 6472377
; GENERAL INFORMATION:
; APPLICANT: Frittz, Christian
; APPLICANT: Youngman, Philip
; TITLE OF INVENTION: USE OF YNES, ESSENTIAL BACTERIAL GENES AND POLYPEPTIDES
; FILE REFERENCE: 06286-C90001
; CURRENT APPLICATION NUMBER: US/09/163,445B
; CURRENT FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: US 60/070,116
; PRIOR FILING DATE: 1997-12-31
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-163-445B-1

Query Match      3.5%: Score 7; DB 4; Length 213;
Best Local Similarity 100.0%: Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AVATSAG 36
Db 105 AVATSAG 111

RESULT 6
US-09-393-634-55
; Sequence 55, Application US/09393634
; Patent No. 6558910
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Ryba, Nick
; APPLICANT: Mueller, Ken
; APPLICANT: Hoon, Mark
; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: SP, a No. 6558910el Family of Taste Receptors
; FILE REFERENCE: 02307E-098000S
; CURRENT APPLICATION NUMBER: US/09/393,634
; CURRENT FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 55
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human GR11
; NAME/KEY: MOD_RES
; LOCATION: (1)..(245)
; OTHER INFORMATION: Xaa any amino acid
US-09-393-634-55

Query Match      3.5%: Score 7; DB 4; Length 245;
Best Local Similarity 100.0%: Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 SLIDFIL 162
Db 42 SLIDFIL 43

RESULT 7
US-09-393-634-56
; Sequence 56, Application US/09393634
; Patent No. 6558910
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Zuker, Charles S.
/ APPLICANT: Adler, Jon Elliot
/ APPLICANT: Ryba, Nick
/ APPLICANT: Mueller, Ken
/ APPLICANT: Hoon, Mark
/ APPLICANT: The Regents of the University of California
/ APPLICANT: The Government of the United States of America
/ APPLICANT: as represented by the Secretary of the
/ APPLICANT: Department of Health and Human Services
/ TITLE OF INVENTION: SF, a No. 6558903el Family of Taste Receptors
/ FILE REFERENCE: C2307E-C98003US
/ CURRENT APPLICATION NUMBER: US/09/333,634
/ CURRENT FILING DATE: 1999-09-10
/ NUMBER OF SEQ ID NOS: 92
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 56
/ LENGTH: 315
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: MOD_RES
/ LOCATION: (1)..(315)
/ OTHER INFORMATION: xaa - any amino acid
US-09-393-634-56

Query Match          3.5%; Score 7; DB 4; Length 315;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 SLIDFIL 162
DB 42 SLIDFIL 48

RESULT 8
US-09-345-473E-45
/ Sequence 45, Application US/09345473E
/ Patent No. 6558903
/ GENERAL INFORMATION:
/ APPLICANT: Hodges, Martin
/ TITLE OF INVENTION: No. 6558903el Kinases and Uses Thereof
/ FILE REFERENCE: 35800/183781
/ CURRENT APPLICATION NUMBER: US/09/345,473E
/ CURRENT FILING DATE: 1999-06-30
/ NUMBER OF SEQ ID NOS: 62
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 45
/ LENGTH: 451
/ TYPE: PRT
/ ORGANISM: Mus musculus
US-09-345-473E-45

Query Match          3.5%; Score 7; DB 4; Length 451;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 ALSFLG 127
DB 416 ALSFLG 422

RESULT 9
US-09-328-352-6888
/ Sequence 6888, Application US/09328352
/ Patent No. 6562958
/ GENERAL INFORMATION:
/ APPLICANT: Gary L. Breton et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ADINETOBACTER
/ FILE REFERENCE: GTC99-03PA
/ CURRENT APPLICATION NUMBER: US/09/328,352

/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMON/
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ CURRENT FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 62/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 29267
/ LENGTH: 483
/ TYPE: PRT
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29267

Query Match          3.5%; Score 7; DB 4; Length 483;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 PALGLK 145
DB 438 PALGLK 444

RESULT 11
US-09-345-473E-44
/ Sequence 44, Application US/09345473E
/ Patent No. 6558903
/ GENERAL INFORMATION:
/ APPLICANT: Hodges, Martin
/ TITLE OF INVENTION: No. 6558903el Kinases and Uses Thereof
/ FILE REFERENCE: 35800/183781
/ CURRENT APPLICATION NUMBER: US/09/345,473E
/ CURRENT FILING DATE: 1999-06-30
/ NUMBER OF SEQ ID NOS: 62
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 44
/ LENGTH: 617
/ TYPE: PRT
/ ORGANISM: Mus musculus
US-09-345-473E-44

Query Match          3.5%; Score 7; DB 4; Length 617;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 ALSFLG 127
DB 582 ALSFLG 588
```

RESULT 12

US-09-252-991A-26617
 ; Sequence 26617, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252.991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 26617
 ; LENGTH: 643
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-26617

Query Match 3.5%, Score 7; DB 4; Length 643;
 Best Local Similarity 100.0%; Pred. No. 92;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GPSAPEA 9

|||||

Db 1 GPSAPEA 7

RESULT 13

US-09-063-950-2
 ; Sequence 2, Application US/09063950C
 ; Patent No. 6225085
 ; GENERAL INFORMATION:
 ; APPLICANT: Hotzman, Douglas A.
 ; TITLE OF INVENTION: NOVEL LRSG PROTEIN AND NUCLEIC ACID MOLECULES AND USES
 ; FILE REFERENCE: MEI-019
 ; CURRENT APPLICATION NUMBER: US/09/063.950C
 ; CURRENT FILING DATE: 1998-04-21
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 673
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-063-950-2

Query Match 3.5%, Score 7; DB 3; Length 673;
 Best Local Similarity 100.0%; Pred. No. 96;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 ITNETFR 196

|||||

Db 115 ITNETFR 121

RESULT 14

US-09-996-243-52
 ; Sequence 52, Application US/09996243
 ; Patent No. 6478825
 ; GENERAL INFORMATION:
 ; APPLICANT: Ashkenazi, Avi J.
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Eric
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Fong, Sherman

APPLICANT: Gerber, Hanspeter
 APPLICANT: Gerritsen, Mary E.
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Grimaldi, J. Christopher
 APPLICANT: Gurney, Austin I.
 APPLICANT: Kljavin, Ivar J.
 APPLICANT: Napier, Mary A.
 APPLICANT: Pan, James
 APPLICANT: Paoni, Nicholas F.
 APPLICANT: Roy, Margaret Ann
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Watanabe, Colin K.
 APPLICANT: Williams, P. Mickey
 APPLICANT: Wood, William L.
 APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 TITLE OF INVENTION: Acids Encoding the Same
 FILE REFERENCE: P2730P1C13
 CURRENT APPLICATION NUMBER: US/09/996.243
 CURRENT FILING DATE: 2002-11-14
 PRIOR APPLICATION NUMBER: 60/049787
 PRIOR FILING DATE: 1997-06-16
 PRIOR APPLICATION NUMBER: 60/062250
 PRIOR FILING DATE: 1997-10-17
 PRIOR APPLICATION NUMBER: 60/065186
 PRIOR FILING DATE: 1997-11-12
 PRIOR APPLICATION NUMBER: 60/065311
 PRIOR FILING DATE: 1997-11-13
 PRIOR APPLICATION NUMBER: 60/066770
 PRIOR FILING DATE: 1997-11-24
 PRIOR APPLICATION NUMBER: 60/075945
 PRIOR FILING DATE: 1998-02-25
 PRIOR APPLICATION NUMBER: 60/078910
 PRIOR FILING DATE: 1998-03-20
 PRIOR APPLICATION NUMBER: 60/083322
 PRIOR FILING DATE: 1998-04-28
 PRIOR APPLICATION NUMBER: 60/084600
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/087106
 PRIOR FILING DATE: 1998-05-28
 PRIOR APPLICATION NUMBER: 60/087607
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087609
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087759
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087827
 PRIOR FILING DATE: 1998-06-03
 PRIOR APPLICATION NUMBER: 60/088021
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088025
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088026
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088028
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088029
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088030
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088033
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088326
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088167
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088202
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088212
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088217

;
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088875
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088905
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089607
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24

;
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 3.5%; Score 7; DB 4; Length 673;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 :TNETER 1%
DB 115 ITNETER 121

RESULT 15

US-59-252-99A-29401
; Sequence 29401, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29401
; LENGTH: 707
; TYPE: PRT

: ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29401

Query Match	3.5%	Score 7	DB 4	Length 707
Best Local Similarity	100.0%	Pred. No. 1e+02	0	Indels
Matches	7	Conservative	0	Mismatches
			0	Gaps
QY	100	VGFEKPI	106	
DB	36	VGFEKPI	42	

Search completed: September 26, 2003, 17:43:23
Job time : 37.1021 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2003, 17:41:48 : Search time 70.4851 Seconds
(without alignments)
433.624 Million cell: updates/sec

Title: US-99-852-100A-2_COPY 68 259

perfect score: 202
Sequence: 1 PSGPSAPEAVTARLVGVLF.....TETRLSITNEPRKTOINP 202

Scoring table: OLIGO

Searched: 565894 seqs, 151307093 residues
Gapop 60.0 , Gapext 50.0

```

Word size : 0
Total number of hits satisfying chosen parameters: 526894

```

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Listing first 45 summaries

```
Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMP.pcp.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PCB.pcp.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMP.pcp.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMP.pcp.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PCB.pcp.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMP.pcp.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMP.pcp.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMP.pcp.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMP.pcp.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMP.pcp.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMP.pcp.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMP.pcp.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMP.pcp.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	202	100.0	0	269	9	US-09-852-100A-2	Sequence 2, Appl
2	202	100.0	0	269	10	US-09-833-503A-2	Sequence 2, Appl
3	202	100.0	0	269	15	US-10-199-88-2	Sequence 2, Appl
4	8	4.0	0	21	15	US-10-226-956-288	Sequence 288, App
5	8	4.0	0	21	15	US-10-111-088-306	Sequence 306, App
6	8	4.0	0	24	11	US-09-615-914B-28	Sequence 28, Appl
7	8	4.0	0	27	11	US-09-785-802A-8	Sequence 8, Appl
8	8	4.0	0	27	11	US-09-965-876A-2	Sequence 2, Appl
9	8	4.0	0	27	12	US-09-915-914B-27	Sequence 27, Appl
10	8	4.0	0	27	15	US-10-077-555-8	Sequence 8, Appl
11	8	4.0	0	221	10	US-09-833-503A-6	Sequence 6, Appl
12	8	4.0	0	221	11	US-09-992-606A-62	Sequence 62, Appl
13	8	4.0	0	221	11	US-09-924-140-62	Sequence 82, Appl
14	8	4.0	0	221	11	US-09-974-879-230	Sequence 23, Appl
15	8	4.0	0	221	12	US-09-952-095B-82	Sequence 82, Appl

16	8	4.0	221	12	US-10-154-678-82	Sequence 82, Appl
17	8	4.0	221	12	US-09-999-570-82	Sequence 82, Appl
18	8	4.0	221	12	US-10-000-489-82	Sequence 82, Appl
19	8	4.0	221	15	US-10-000-986-82	Sequence 82, Appl
20	8	4.0	221	15	US-10-199-881-6	Sequence 6, Appl
21	8	4.0	222	11	US-09-305-736-230	Sequence 230, Appl
22	8	4.0	230	11	US-09-796-753-50	Sequence 50, Appl
23	8	4.0	247	11	US-09-796-753-48	Sequence 48, Appl
24	8	4.0	358	12	US-10-174-794-11	Sequence 11, Appl
25	7	3.5	114	15	US-10-154-251-13	Sequence 13, Appl
26	7	3.5	122	9	US-09-867-550-80	Sequence 80, Appl
27	7	3.5	173	11	US-09-946-374-332	Sequence 332, Appl
28	7	3.5	173	12	US-10-015-367A-332	Sequence 332, Appl
29	7	3.5	173	12	US-10-006-130A-332	Sequence 332, Appl
30	7	3.5	173	12	US-10-006-172A-332	Sequence 332, Appl
31	7	3.5	173	12	US-10-015-392A-332	Sequence 332, Appl
32	7	3.5	173	12	US-10-017-253A-332	Sequence 332, Appl
33	7	3.5	173	12	US-10-017-306A-332	Sequence 332, Appl
34	7	3.5	173	15	US-10-006-856A-332	Sequence 332, Appl
35	7	3.5	173	15	US-10-006-818A-332	Sequence 332, Appl
36	7	3.5	173	15	US-10-015-363A-332	Sequence 332, Appl
37	7	3.5	173	15	US-10-015-869A-332	Sequence 332, Appl
38	7	3.5	173	15	US-10-012-121A-332	Sequence 332, Appl
39	7	3.5	173	15	US-10-006-116A-332	Sequence 332, Appl
40	7	3.5	173	15	US-10-006-117A-332	Sequence 332, Appl
41	7	3.5	173	15	US-10-017-527A-332	Sequence 332, Appl
42	7	3.5	173	15	US-10-017-913A-332	Sequence 332, Appl
43	7	3.5	173	15	US-10-007-194A-332	Sequence 332, Appl
44	7	3.5	173	15	US-10-013-430A-332	Sequence 332, Appl
45	7	3.5	173	15	US-10-011-671A-332	Sequence 332, Appl

ALIGNMENT'S

```

RESULT 1
US-09-852-100A-2
: Sequence 2, Application US/09852-100A
: Patent No. US2002058267A1
: GENERAL INFORMATION:
: APPLICANT: American Home Products
: TITLE OF INVENTION: Beta-amyloid Peptide-B
: TITLE OF INVENTION: Same
: FILE REFERENCE: AHP981261p2
: CURRENT APPLICATION NUMBER: US/09/852,100A
: CURRENT FILING DATE: 2001-05-09
: PRIOR APPLICATION NUMBER: US/09/172,990
: PRIOR FILING DATE: 1998-10-14
: PRIOR APPLICATION NUMBER: US/60/104,104
: PRIOR FILING DATE: 1998-10-13
: PRIOR APPLICATION NUMBER: PTC/US99/21621
: PRIOR FILING DATE: 1999-10-13
: PRIOR APPLICATION NUMBER: US/09/060,609
: PRIOR FILING DATE: 1998-04-15
: PRIOR APPLICATION NUMBER: US/60/64,583
: PRIOR FILING DATE: 1997-04-16
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: Patentin version 3.0
: SEQ ID NO 2
: LENGTH: 269
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-852-100A-2

```

```

Query Match: 100.0%; Score 202; DB 9; Length 269;
Best local Similarity 100.0%; Pred. No. 28-190;
Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGGSPAEAVTARLVGLVFWFVTTGPMGAVATSGAGEESLKCEDLVQGYICKDPKIND 60
|||||
DB 68 PGGSPAEAVTARLVGLVFWFVTTGPMGAVATSGAGEESLKCEDLVQGYICKDPKIND 127
|||||

QY 61 ATQEPVNCNTIAIVSCFFPAPNITCKOSSGNETHFTGNLWGFEEKPISCRNNVNGYSKKVAV 120

```



```
; APPLICANT: Chen, Yih-Tai
; TITLE OF INVENTION: NO. US20030104479A1el Fusion Proteins And Assays For Molecular Bi
; FILE REFERENCE: 01-1022-US
; CURRENT APPLICATION NUMBER: US/10/211,088
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/309,395
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/341,589
; PRIOR FILING DATE: 2001-12-13
; NUMBER OF SEQ ID NOS: 366
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 306
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Protein-derived transport peptide
US-10-211-088-306

Query Match          4.0%; Score 8; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 LFLGWLGA 131
DB 3 LFLGWLGA 10

RESULT 6
US-09-915-914B-28
; Sequence 28, Application US/09015914B
; Publication No. US20030119725A1
; GENERAL INFORMATION:
; APPLICANT: Divita, Gilles
; APPLICANT: Fernandez, Joseph
; APPLICANT: Heitz, Frederic
; APPLICANT: Morris, May
; APPLICANT: Mery, Jean
; APPLICANT: Archdeacon, John
; APPLICANT: Horndorp, Kyle
; TITLE OF INVENTION: PEPTIDE-MEDIATED DELIVERY OF MOLECULES INTO CELLS
; FILE REFERENCE: AM-00105.P.1.1
; CURRENT APPLICATION NUMBER: US/09/9:5,914B
; CURRENT FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: US 60/221,932
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic sequence
US-09-915-914B-28

Query Match          4.0%; Score 8; DB 11; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 LFLGWLGA 131
DB 3 LFLGWLGA 10

RESULT 7
US-09-785-802A-8
; Sequence 8, Application US/09785802A
; Patent No. US20020151004A1
; GENERAL INFORMATION:
; APPLICANT: Craig, Roger
; TITLE OF INVENTION: DELIVERY VEHICLES AND METHODS FOR USING THE SAME
; FILE REFERENCE: 11067/2035
```

```
; CURRENT APPLICATION NUMBER: US/09/785,802A
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 09/748,06
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/748,789
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-785-802A-8

Query Match          4.0%; Score 8; DB 10; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 LFLGWLGA 131
DB 3 LFLGWLGA 10

RESULT 8
US-09-965-876A-2
; Sequence 2, Application US/09965876A
; Publication No. US2003036243A1
; GENERAL INFORMATION:
; APPLICANT: Busa, William B
; TITLE OF INVENTION: Methods and Reagents for Live-cell Gene Expression Quantifica
; FILE REFERENCE: 00-789-A
; CURRENT APPLICATION NUMBER: US/09/965,876A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/236,407
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 27
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-09-965-876A-2

Query Match          4.0%; Score 8; DB 11; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 LFLGWLGA 131
DB 3 LFLGWLGA 10

RESULT 9
US-09-915-914B-27
; Sequence 27, Application US/09915914B
; Publication No. US20030119725A1
; GENERAL INFORMATION:
; APPLICANT: Divita, Gilles
; APPLICANT: Fernandez, Joseph
; APPLICANT: Heitz, Frederic
; APPLICANT: Morris, May
; APPLICANT: Mery, Jean
; APPLICANT: Archdeacon, John
; APPLICANT: Horndorp, Kyle
; TITLE OF INVENTION: PEPTIDE-MEDIATED DELIVERY OF MOLECULES INTO CELLS
; FILE REFERENCE: AM-00105.P.1.1
; CURRENT APPLICATION NUMBER: US/09/915,914B
; CURRENT FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: US 60/221,932
; PRIOR FILING DATE: 2000-07-31
```

1 NUMBER OF SEQ ID NOS: 43
2 SOFTWARE: PatentIn version 3.1
3 SEQ ID NO 27
4 LENGTH: 27
5 TYPE: PRT
6 ORGANISM: Artificial Sequence
7 FEATURE:
8 OTHER INFORMATION: Synthetic sequence
9
10 US-09-915-914B-27

Query Match 4.0% Score 8; DB 11; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 LFLGWLGA 131
|||||
DB 3 LFLGWLGA 10

RESULT 10

US-10-077-555-8
1 Sequence 8, Application US/10077555
2 Publication No. US2003007289A1
3 GENERAL INFORMATION:
4 APPLICANT: Wang, Rong-fu
5 TITLE OF INVENTION: Use of Cell-Penetrating Peptides to Generate Antitumor Immunity
6 FILE REFERENCE: P02373US/10200806
7 CURRENT APPLICATION NUMBER: US/10/077.555
8 CURRENT FILING DATE: 2002-02-15
9 PRIOR APPLICATION NUMBER: US 60/268,667
10 PRIOR FILING DATE: 2001-02-15
11 NUMBER OF SEQ ID NOS: 14
12 SOFTWARE: PatentIn version 3.1
13 SEQ ID NO 8
14 LENGTH: 27
15 TYPE: PRT
16 ORGANISM: Artificial Sequence
17 FEATURE:
18 OTHER INFORMATION: Synthetic Peptide
19
20 US-10-077-555-8

Query Match 4.0% Score 8; DB 15; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 LFLGWLGA 131
|||||
DB 3 LFLGWLGA 10

RESULT 11

US-09-833-503A-6
1 Sequence 6, Application US/09833503A
2 Patent No. US20020146760A1
3 GENERAL INFORMATION:
4 APPLICANT: Ozenberger, Bradley A
5 APPLICANT: Kajkowski, Eileen M
6 APPLICANT: Lo, Ching-Hsiung F
7 APPLICANT: American Home Products Corporation
8 TITLE OF INVENTION: No. US20020146760A1el G-Protein-Coupled Receptor-Like Proteins and
9 TITLE OF INVENTION: Polynucleotides Encoded By Them, and Methods of Using
10 TITLE OF INVENTION: Same
11 FILE REFERENCE: AHP98165-00PCT
12 CURRENT APPLICATION NUMBER: US/09/833.503A
13 CURRENT FILING DATE: 2000-10-13
14 PRIOR APPLICATION NUMBER: 60/104,104
15 PRIOR FILING DATE: 1998-10-13
16 NUMBER OF SEQ ID NOS: 6
17 SOFTWARE: PatentIn Ver. 2.1
18 SEQ ID NO 6
19 LENGTH: 221
20 TYPE: PRT
21 ORGANISM: Homo sapiens

US-09-833-503A-6

Query Match 4.0% Score 8; DB 10; Length 221;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GADREYLG 137
|||||
DB 174 GADREYLG 181

RESULT 12

US-09-992-600A-82
1 Sequence 82, Application US/09992600A
2 Publication No. US20030027161A1
3 GENERAL INFORMATION:
4 APPLICANT: Benjanin, Stephanie
5 APPLICANT: Tanaka, Hiroaki
6 TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
7 FILE REFERENCE: 91.US4.DIV
8 CURRENT APPLICATION NUMBER: US/09/992.600A
9 CURRENT FILING DATE: 2001-11-13
10 PRIOR APPLICATION NUMBER: US 09/924,340
11 PRIOR FILING DATE: 2001-08-06
12 PRIOR APPLICATION NUMBER: PCT/IB01/01715
13 PRIOR FILING DATE: 2001-08-06
14 PRIOR APPLICATION NUMBER: US 60/305,456
15 PRIOR FILING DATE: 2001-07-13
16 PRIOR APPLICATION NUMBER: US 60/302,277
17 PRIOR FILING DATE: 2001-06-29
18 PRIOR APPLICATION NUMBER: US 60/298,698
19 PRIOR FILING DATE: 2001-06-15
20 PRIOR APPLICATION NUMBER: US 60/293,574
21 PRIOR FILING DATE: 2001-05-25
22 NUMBER OF SEQ ID NOS: 114
23 SOFTWARE: JPatent
24 SEQ ID NO 82
25 LENGTH: 221
26 TYPE: PRT
27 ORGANISM: Homo sapiens
28 FEATURE:
29 NAME/KEY: SIGNAL
30 LOCATION: 1..32
31
32 US-09-992-600A-82

Query Match 4.0% Score 8; DB 11; Length 221;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GADREYLG 137
|||||
DB 174 GADREYLG 181

RESULT 13

US-09-924-340-82
1 Sequence 82, Application US/09924340
2 Publication No. US20030027248A1
3 GENERAL INFORMATION:
4 APPLICANT: Benjanin, Stephanie
5 APPLICANT: Tanaka, Hiroaki
6 TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
7 FILE REFERENCE: 91.US2.REG
8 CURRENT APPLICATION NUMBER: US/09/924.340
9 CURRENT FILING DATE: 2001-08-06
10 PRIOR APPLICATION NUMBER: US 60/305,456
11 PRIOR FILING DATE: 2001-07-13
12 PRIOR APPLICATION NUMBER: US 60/302,277
13 PRIOR FILING DATE: 2001-06-29
14 PRIOR APPLICATION NUMBER: US 60/298,698
15 PRIOR FILING DATE: 2001-06-15
16 PRIOR APPLICATION NUMBER: US 60/293,574
17 PRIOR FILING DATE: 2001-05-25

```

; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 82
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: 1..32
US-09-924-340-82

```

```

Query Match          4.0%  Score 8;  DB 11;  Length 221;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches      8;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

QY      130 GADREYLG 137
      |||||
DB      174 GADREYLG 181

```

RESULT 14

```

; Sequence 230, Application US/0974879
; Publication No. US2003002803A1
; GENERAL INFORMATION:
; APPLICANT: Kosen et al.
; TITLE OF INVENTION: 125 Human Secreted Proteins
; FILE REFERENCE: P2029P2
; CURRENT APPLICATION NUMBER: US/09/974,879
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/233,893
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/818,683
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 09/305,736
; PRIOR FILING DATE: 1999-05-05
; PRIOR APPLICATION NUMBER: PCT/US98/23435
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: US 60/064,911
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,912
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,983
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,900
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,988
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,987
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,908
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,984
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,985
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/066,034
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,100
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/065,089
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,095
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,090
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 611
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 230
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

```

```

; NAME/KEY: SITE
; LOCATION: (184)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-974-879-230

```

```

Query Match          4.0%  Score 8;  DB 12;  Length 221;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches      8;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

QY      130 GADREYLG 137
      |||||
DB      174 GADREYLG 181

```

RESULT 15

```

; US-09-992-095B-82
; Sequence 82, Application US/09992095B
; Publication No. US20030157485A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephanie
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91,US5, DIV
; CURRENT APPLICATION NUMBER: US/09/992,095B
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: US 60/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 82
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: 1..32
US-09-992-095B-82

```

```

Query Match          4.0%  Score 8;  DB 12;  Length 221;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches      8;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

QY      130 GADREYLG 137
      |||||
DB      174 GADREYLG 181

```

```

Search completed: September 26, 2003, 17:55:12
Job time : 71.4851 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2003, 17:29:58 ; Search time 44.6979 Seconds
(without alignments)
434.608 Million cell updates/sec

Title: US-09-852-100A-2_COPY_68_269

Perfect score: 202
Sequence: 1 PSGPSAPEAVTARLVGVLMF.....TRLTRLSITNTRFKVQLNP 202

Scoring table: OLLGO
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs. 95168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR.76.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	8	4.0	358	2	S23383	protein kinase (EC 2.7.1.37)
2	8	4.0	376	2	S22745	serine/threonine protein kinase
3	8	4.0	548	2	T16682	hypothetical prote
4	8	4.0	753	2	T28797	hypothetical prote
5	7	3.5	97	2	B97574	hypothetical prote
6	7	3.5	102	2	H70898	probable PF protei
7	7	3.5	141	2	D75155	acetyl transferase
8	7	3.5	155	2	T48956	hypothetical prote
9	7	3.5	173	2	G71916	inorganic pyrophos
10	7	3.5	173	2	D64597	inorganic pyrophos
11	7	3.5	185	2	T20157	hypothetical prote
12	7	3.5	198	2	AB1598	conserved hypotet
13	7	3.5	198	2	AD1235	B. subtilis YncS p
14	7	3.5	202	2	A89911	conserved hypotet
15	7	3.5	204	2	F81445	type IV pilus asse
16	7	3.5	213	2	B86747	conserved hypotet
17	7	3.5	213	2	C95098	conserved hypotet
18	7	3.5	213	2	C97966	conserved hypotet
19	7	3.5	215	2	C69937	conserved hypotet
20	7	3.5	219	2	B75533	hypothetical prote
21	7	3.5	248	2	F71538	probable oxoacyl (
22	7	3.5	273	2	C71369	probable sugar ABC
23	7	3.5	293	2	D70106	conserved hypotet
24	7	3.5	307	2	T22128	hypothetical prote
25	7	3.5	317	2	A71237	hypothetical prote
26	7	3.5	343	2	B72507	hypothetical prote
27	7	3.5	349	2	B81736	probable 3'(2'),5'
28	7	3.5	384	2	AB0147	probable ABC trans
29	7	3.5	389	2	G61062	trans-sulfuration

30 7 3.5 415 2 C83544 probable MFS trans
31 7 3.5 418 2 D86938 probable para-amin
32 7 3.5 419 1 J00346 ubiquinol-cytochrc
33 7 3.5 442 2 B75634 hypothetical prote
34 7 3.5 451 1 JE0240 LIM kinase (EC 2.7
35 7 3.5 458 2 H70602 probable pabb prot
36 7 3.5 466 2 T44650 capsular polysacch
37 7 3.5 492 2 G75389 NADH2 dehydrogenas
38 7 3.5 495 2 A95984 probable xanthine
39 7 3.5 506 2 T12819 hypothetical prote
40 7 3.5 510 2 B71017 hypothetical prote
41 7 3.5 511 1 IKECB8 colicin B - Escher
42 7 3.5 540 2 C84744 probable pp19-lik
43 7 3.5 547 2 H75632 Na(+)-linked D-ala
44 7 3.5 575 2 T43400 myo-inositol trans
45 7 3.5 583 2 T25690 hypothetical prote

ALIGNMENTS

RESULT 1

S23383 protein kinase (EC 2.7.1.37) cdc2-related KKIALRE - human

C:Species: Homo sapiens (man)

C>Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 18-Jun-1999

C:Accession: S23383; S22744

R:Myerson, M.; Enders, G.H.; Wu, C.L.; Su, L.K.; Gorka, C.; Nelson, C.; Harlow, E.

EMBO J. 11, 2909-2917, 1992

A:Title: A family of human cdc2-related protein kinases.

A:Reference number: S23382; MUID:92347325; PMID:1639063

A:Accession: S23383

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-358 <MEY>

A:Cross-references: EMBL:X66358; NID:q36614; PDB:CAA47002.1; PID:q36615

C:Superfamily: Kinase-related transforming protein; protein kinase homology

C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase

F:3-278/Domain: protein kinase homology <KIN>

F:11-19/Region: protein kinase ATP-binding motif

F:34,52,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match: 4.0%; Score 8; DB 2; Length 358;
Best local similarity 100.0%; Pred. No. 3.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 138 YPALGLLK 145

DB 259 YPALGLLK 266

RESULT 2

S22745 serine/threonine protein kinase KKIALRE (EC 2.7.1.-) - human

C:Species: Homo sapiens (man)

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 19-Dec-1997

C:Accession: S22745

R:Myerson, M.L.

submitted to the EMBL Data Library, May 1992

A:Reference number: S22743

A:Accession: S22745

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-376 <MEY>

A:Cross-references: EMBL:X66359

C:Genetics: 152/3; 170/3

C:Superfamily: Kinase-related transforming protein; protein kinase homology

C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase

F:3-296/Domain: protein kinase homology <KIN>

F:11-19/Region: protein kinase ATP-binding motif

Query Match 4.0%; Score 8; DB 2; Length 376;

```

Best Local Similarity 100.0%; Pred. No. 3/2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 YPALGLLK 145
    |||||
Db 277 YPALGLLK 284

RESULT 3
T16082
hypothetical protein F16H11.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 18-Feb-2000
C:Accession: T16082
R:Wo, X.
submitted to the EMBL Data Library, April 1996
A:Description: The sequence of C. elegans cosmid F16H11.
A:Reference number: Z18458
A:Accession: T16082
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-548 <WUX>
A:Cross-references: EMBL:U55376; NID:q1280130; PID:q1280133; PIDN:AAA98005.1; GSPDB:GN000
A:Experimental source: strain Bristol N2; clone F16H11
C:Genetics:
A:Gene: CESP:F16H11.5
A:Map position: X
A:Introns: 71/1: 93/1: 187/3: 281/3: 316/3: 420/2: 460/1: 523/3: 545/3

Query Match 4.0%; Score 8; DB 2: Length 548;
Best Local Similarity 100.0%; Pred. No. 4/6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 KOSSGNET 93
    |||||
Db 109 KOSSGNET 116

RESULT 4
T28787
hypothetical protein C41D11.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T28787
R:Gattung, S.; Maggi, L.
submitted to the EMBL Data Library, May 1997
A:Description: The sequence of C. elegans cosmid C41D11.
A:Reference number: Z20522
A:Accession: T28787
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-753 <GAT>
A:Cross-references: EMBL:AF003740; PIDN:AA048141.1; GSPDB:GN00019; CESP:C41D11.5
A:Experimental source: strain Bristol N2; clone C41D11
C:Genetics:
A:Gene: CESP:C41D11.5
A:Map position: 1
A:Introns: 53/2: 81/3: 117/1: 256/3: 274/2: 357/3: 443/2: 485/3: 544/3: 585/3: 637/2

Query Match 4.0%; Score 8; DB 2: Length 753;
Best Local Similarity 100.0%; Pred. No. 6/1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GADRFYLG 137
    |||||
Db 398 GADRFYLG 405

RESULT 5
B97574
hypothetical protein AGR_C_3268 [imported] - Agrobacterium tumefaciens (strain C56, Cure
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002

C:Accession: B97574
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Oudollo, B.; Goic
A.: Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Iappas, C.; Markelz,
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: B97574
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-97 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK87547.1; PID:q15156883; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_3268
A:Map position: circular chromosome

Query Match 3.5%; Score 7; DB 2: Length 97;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 EAVTARL 14
    |||||
Db 47 EAVTARL 53

RESULT 6
H70898
Probable PE protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: H70898
R:Gale, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gorde
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Koltoyc
Raandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete gen
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: H70898
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-102 <COL>
A:Cross-references: GB:280108; GB:AL123456; NID:q3256012; PIDN:CAB02191.1; PID:e2655
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: PE

Query Match 3.5%; Score 7; DB 2: Length 102;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 EAVTARL 14
    |||||
L6 18 EAVTARL 24

RESULT 7
D75155
acetyl transferase PAB2081 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: D75155
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosom
A:Reference number: A75001
A:Accession: D75155
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-141 <KAW>
A:Cross-references: GB:AJ248284; GB:AL096836; NID:q5457730; PIDN:CAB49323.1; PID:el5
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB2081

```



```

Query Match          3.5%  Score 7:  DB 2:  Length 141;
Best Local Similarity 100.0%  Pred. No. 15;
Matches 7:  Conservative 0;  Mismatches 0;  Gaps 0;

Qy  149 VGCGIG 155
Db   42 VGCGIG 48

RESULT 8
T48956
hypothetical protein T15B3.160 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 13-Aug-2000
C:Accession: T48956
R:Jordan, N.; Bangert, S.; Wiedemann, R.; Voss, H.; Unselid, M.; Mewes, H.W.; Radd, S.;
submitted to the Protein Sequence Database, April 2000
A:Reference number: 225009
A:Accession: T48956
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-159 <OR>
A:Cross-references: EMBL:AL163975; GSPDB:GN00061; ATSP:T15B3.160
A:Experimental source: cultivar Columbia; BAC clone T15B3
C:Genetics:
A:Gene: ATSP:T15B3.160
A:Map position: 3
A:Introns: 102/3; 115/1; 143/1
C:Superfamily: Arabidopsis thaliana hypothetical protein T15B3.160

Query Match          3.5%  Score 7:  DB 2:  Length 159;
Best Local Similarity 100.0%  Pred. No. 17;
Matches 7:  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy  121 ALSIFLG 127
Db   94 ALSIFLG 100

RESULT 9
G71916
inorganic pyrophosphatase - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 22-Jun-1999
C:Accession: G71916
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, H.L.; Brown, E.D.; Doig, P.C.; Smith, P.R.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Voytes, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: G71916
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-173 <ARN>
A:Cross-references: GB:AE001489; GB:AE001439; NID:94155102; PIDN:AA066-46.1; PID:94.551-
A:Experimental source: strain J99
C:Genetics:
A:Gene: ppa
C:Superfamily: inorganic pyrophosphatase

Query Match          3.5%  Score 7:  DB 2:  Length 173;
Best Local Similarity 100.0%  Pred. No. 19;
Matches 7:  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy  12 ARLVGL 18
Db   86 ARLVGL 92

RESULT 10
D64597
inorganic pyrophosphatase - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C>Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 22-Jun-1999
C:Accession: D64597
R:Tomb, J.F.; White, O.; Keriavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann,
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, T.; McI
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Vattthei
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: D64597
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-173 <OM>
A:Cross-references: GB:AE000576; GB:AE000511; NID:92313736; PIDN:AA07684.1; IID:92;
C:Superfamily: inorganic pyrophosphatase

Query Match          3.5%  Score 7:  DB 2:  Length 173;
Best Local Similarity 100.0%  Pred. No. 19;
Matches 7:  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy  12 ARLVGL 18
Db   86 ARLVGL 92

RESULT 11
T20257
hypothetical protein F58H1.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 23-Sep-2002
C:Accession: T20157; T22958
R:Holt, R.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z19230
A:Accession: T20157
A>Status: preliminary; translated from GB/EMBL/DD83
A:Molecule type: DNA
A:Residues: 1-185 <WIL>
A:Cross-references: EMBL:Z75012; PIDN:CAB01419.1; GSPDB:GN000023; CESP:F58H1.7
A:Experimental source: clone C52E4
R:Harris, B.
submitted to the EMBL Data Library, July 1996
A:Reference number: Z19642
A:Accession: T22958
A>Status: preliminary; translated from GB/EMBL/DD83
A:Molecule type: DNA
A:Residues: 1-185 <W12>
A:Cross-references: EMBL:Z75954; PIDN:CAB00110.1; GSPDB:GN000023; CESP:F58H1.7
C:Genetics:
A:Gene: CFSP:F58H1.7
A:Map position: 5
A:Introns: 33/1; 77/1; 103/3; 128/1
C:Superfamily: LR11 protein; laminin-type EGF-like homology; LDL receptor ligand-bu
F:37-74/Domain: LDL receptor ligand-binding repeat homology <LD1>

Query Match          3.5%  Score 7:  DB 2:  Length 185;
Best Local Similarity 100.0%  Pred. No. 20;
Matches 7:  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy  117 KVAVALS 123
Db  121 KVAVALS 127

RESULT 12
AB1598
Conserved hypothetical protein, B. subtilis YneS protein homolog lin1323 [imported]
C:Species: Listeria innocua
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AB1598

```

R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, J.; Dussurget, C.; Entian, K.D.; Fsihi, H. J. Science 254, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkai, G.; Madueno, E.; Maitournam, A.; Maier, C.; Schluteter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, C.; Schluteter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, C.

A:Title: Comparative genomics of *Listeria* species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AB1598

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-198 <GIA>

A:Cross-references: GB:AL592022; PIDN:GAC96554.1; PID:g16413790; GSPDB:GN00178

A:Experimental source: strain Clp11262

C:Genetics:

A:Gene: lml1323

C:Superfamily: Escherichia coli ygiH protein

Query Match 3.5% Score 7: DB 2: Length 198;

Best Local Similarity 100.0%; Pred. No. 21;

Matches 7: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AVATSAG 36
|||||||

DB 109 AVATSAG 115

RESULT 13

AD1235

B. subtilis Ynes protein homolog lml284 [imported] - *Listeria monocytogenes* (strain EGD-8)

C:Species: *Listeria monocytogenes*

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C:Accession: AD1235

R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, J.; Dussurget, C.; Entian, K.D.; Fsihi, H. J. Science 254, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkai, G.; Madueno, E.; Maitournam, A.; Maier, C.; Schluteter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, C.; Schluteter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, C.

A:Title: Comparative genomics of *Listeria* species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AD1235

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-198 <GIA>

A:Cross-references: GB:NC_003210; PIDN:GAC99362.1; PID:g164.0700; GSPDB:GN00177

A:Experimental source: strain EGD-e

C:Genetics:

A:Gene: lml284

C:Superfamily: Escherichia coli ygiH protein

Query Match 3.5% Score 7: DB 2: Length 198;

Best Local Similarity 100.0%; Pred. No. 21;

Matches 7: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AVATSAG 36
|||||||

DB 109 AVATSAG 115

RESULT 14

A89911

Conserved hypothetical protein Sall87 [imported] - *Staphylococcus aureus* (strain N315)

C:Species: *Staphylococcus aureus*

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C:Accession: A89911

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cuatrecasas, J. J. Lancet 357, 1225-1240, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkai, G.; Madueno, E.; Maitournam, A.; Maier, C.; Schluteter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, C.; Schluteter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, C.

A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: A89911

Result No.	Score	Query			ID	Description
		Match	Length	DB		
1	8	4.0	358	1	KK1A_HUMAN	Q00352 homo sapien
2	7	3.5	102	1	YD86_MYCTU	P71656 mycobacteri
3	7	3.5	168	1	YPAE_STRGC	Q9X972 streptococ
4	7	3.5	173	1	IPXR_HELPJ	Q9Z115 helicobacte
5	7	3.5	173	1	IPXR_HELPJ	P56155 helicobacte
6	7	3.5	198	1	YD84_LISMO	Q8V713 listeria in
7	7	3.5	198	1	YD23_LISIN	Q9ZC68 listeria in
8	7	3.5	202	1	YV35_STAEP	P59253 streptococ
9	7	3.5	202	1	YV35_STAEP	Q99uc5 streptococ
10	7	3.5	212	1	YC11_STRMJ	P59255 streptococ
11	7	3.5	212	1	YC10_STRAJ	P59254 streptococ
12	7	3.5	213	1	Y851_STRPN	Q54916 streptococ
13	7	3.5	213	1	Y908_STRPY	Q94070 streptococ
14	7	3.5	213	1	YKAK_LACLA	Q9eqw4 lactococcus
15	7	3.5	215	1	YPTC_BACSU	P43978 bacillus su
16	7	3.5	224	1	CLCH_HUMAN	P56750 homo sapien
17	7	3.5	247	1	FABG_CHLTR	P38004 chlamydia t
18	7	3.5	292	1	SDB2_MOUSE	Q93J20 mus musculin
19	7	3.5	307	1	RB33_CAEEL	Q20365 caenorhabdi
20	7	3.5	322	1	SC13_HUMAN	P55735 homo sapien
21	7	3.5	419	1	CYB_RHOVI	P81378 rhodopsin
22	7	3.5	469	1	CG51_HUMAN	Q9V512 homo sapien
23	7	3.5	510	1	CEAR_ECOLI	P05819 escherichia
24	7	3.5	575	1	ITRI_SCHPO	Q10286 schizosacch
25	7	3.5	638	1	L1K2_RAT	P36670 rattus norv
26	7	3.5	663	1	CNG2_BOVIN	Q03041 bos taurus
27	7	3.5	1039	1	AG43_ECOLI	P39180 escherichia
28	6	3.0	39	1	PSAG_PEA	P20120 pisum sativ
29	6	3.0	58	1	TY13_HUMAN	Q9bz97 homo sapien
30	6	3.0	62	1	PURF_RHOSH	P13402 rhodobacter
31	6	3.0	94	1	PTXB_ECOL6	Q8X7H5 escherichia
32	6	3.0	94	1	PTXB_ECOLI	P37168 escherichia
33	6	3.0	111	1	YG12_BACHD	Q9Kcf9 bacillus ba

Query Match 4.0%; Score 8; DB 1; Length 352;
 Best Local Similarity 100.0%; Pred. No. 2.3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 YPAUGLLK 145
 DB 259 YPAUGLLK 266

RESULT 2
 YD86_MYCTU STANDARD; PRT: 102 AA.
 ID YD86_MYCTU
 AC P71856;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein RV1386 precursor.
 GN RV1386 OR MT1430 OR MTCV2184.03.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 CX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris R.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala R.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies K., Devlin K., Feltwell T., Gentles S., Harlin N., Holroyd S.,
 RA Horsley T., Jagels K., Krohn A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skellon S., Squares R., Rogers J.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrett R.G.,
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544 (1998).
 RN [2]

SEQUENCE FROM N.A.
 STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: Z83108; CAB02191.1;
 DR EMBL: AF207014; AA045695.1;
 DR PIR: H70898; H70898.
 DR TIGR: MT1430;
 DR TubercuList; RV1386;
 DR InterPro: IPR000084; PE_region.
 DR Pfam: PF00934; PE; 1.
 KW Hypothetical protein; Signal; Complete proteome.
 FT SIGNAL 1 31
 FT CHAIN 32 102 HYPOTHETICAL PROTEIN RV1386.
 FT SEQUENCE 102 AA; 9862 MW; 8165F05DBDB9D752 CRC64;
 SQ

Query Match 3.5%; Score 7; DB 1; Length 102;
 Best Local Similarity 100.0%; Pred. No. 6.6;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 8 EAVTARL 14
 DB 18 EAVTARL 24

RESULT 3
 YPAE_STROC STANDARD; PRT: 168 AA.
 ID YPAE_STROC
 AC Q9X972;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein in parE 3' region (ORF2) (Fragment).
 OS Streptococcus gordonii Challis.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 CX NCBI_TaxID=29350;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CHI;
 RA Vriesema A.J., Dankert J., Zaai S.A.;
 RA "Isolation and characterization of promoter regions from Streptococcus
 RT gordonii CHI."
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: BELONGS TO THE UPF0078 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AJ236899; CAB40549.1;
 DR HAMAP: MF_01043; 1.
 DR InterPro: IPR003811; DUF205.
 DR Pfam: PF02660; DUF205; 1.
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 8 28
 FT TRANSMEM 71 91
 FT TRANSMEM 111 132
 FT TRANSMEM 144 164
 FT NON_TER 168 168
 SQ SEQUENCE 168 AA; 17921 MW; A07262BD799A478A CRC64;

Query Match 3.5%; Score 7; DB 1; Length 168;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
 QY 30 AVATSAG 36
 DB 106 AVATSAG 112

RESULT 4
 IPYR_HELPJ STANDARD; PRT: 173 AA.
 ID IPYR_HELPJ
 AC Q92LL5;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Inorganic pyrophosphatase (EC 3.6.1.1) (Pyrophosphate phospho-
 DE hydrolase) (PPase).
 GN PPA OR JHP0564.
 OS Helicobacter pylori J99 (Campylobacter pylori J99).
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
 OC Helicobacteraceae; Helicobacter.
 CX NCBI_TaxID=85963;
 RN [1]

```

RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923582;
RA Alm R.A., Ling E.-S.H., Moir D.I., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild A.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT *Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.;
RL Nature 397:176-180(1999).
CC -!- CATALYTIC ACTIVITY: Diphosphate + H(2)O -> 2 phosphate.
CC -!- COFACTOR: REQUIRES THE PRESENCE OF DIVALENT METAL CATION.
CC MAGNESIUM CONFERS THE HIGHEST ACTIVITY. BINDS 4 DIVALENT CATIONS
CC PER SUBUNIT (BY SIMILARITY).
CC -!- SUBUNIT: Homohexamer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the Ppase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF001489; AAD06146.1; -
DR PIR: G71916; G71916.
DR HSP: P17288; IFAJ.
DR HAMAP: MF_00209; -; 1.
DR InterPro: IPR001596; Pyrophosphatase.
DR Pfam: PF00719; Pyrophosphatase; 1.
DR ProDom: PD002014; Inorg_pphsp; 1.
DR PROSITE: PS00387; PPASE; 1.
KW Hydrolase; Magnesium; Complete proteome.
FT ACT_SITE 28 28 BY SIMILARITY.
SQ SEQUENCE 173 AA; 19258 MW; 87B9B215E6FEBC8 CRC64;

Query Match 3.5%; Score 7; DB 1; Length 173;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 ARLGVGL 18
DB 86 ARLGVGL 92

RESULT 5
ID IPYR_HELPY STANDARD; PRT; 173 AA.
AC P56153;
DT 01-NOV-1997 (Rel. 35; Created)
DI 01-NOV-1997 (Rel. 35; Last sequence update)
DI 28-FEB-2003 (Rel. 41; Last annotation update)
DE Inorganic pyrophosphatase (EC 3.6.1.1) (Pyrophosphate phospho-
DE hydrolase) (PPase).
GN PPA OR HP0620.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kervatage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Melman J.M., Fujii C., Bowman C., Watthey D., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,

```

```

RA Vector J.C.;
RT *The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.;
RI Nature 388:539-547(1997).
CC -!- CATALYTIC ACTIVITY: Diphosphate + H(2)O -> 2 phosphate.
CC -!- COFACTOR: REQUIRES THE PRESENCE OF DIVALENT METAL CATION.
CC MAGNESIUM CONFERS THE HIGHEST ACTIVITY. BINDS 4 DIVALENT CATIONS
CC PER SUBUNIT (BY SIMILARITY).
CC -!- SUBUNIT: Homohexamer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the Ppase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB000576; AAD07684.1; -
DR PIR: D64597; D64597.
DR HSP: P17288; IFAJ.
DR TIGR: HP0620; -
DR HAMAP: MF_00209; -; 1.
DR InterPro: IPR001596; Pyrophosphatase.
DR Pfam: PF00719; Pyrophosphatase; 1.
DR ProDom: PD002014; Inorg_pphsp; 1.
DR PROSITE: PS00387; PPASE; 1.
KW Hydrolase; Magnesium; Complete proteome.
FT ACT_SITE 28 28 BY SIMILARITY.
SQ SEQUENCE 173 AA; 19272 MW; 23A51C665A6ECTF2 CRC64;

Query Match 3.5%; Score 7; DB 1; Length 173;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 ARLGVGL 18
DB 86 ARLGVGL 92

RESULT 6
YC84_L-SMO
ID YC84_LISMOC STANDARD; PRT; 198 AA.
AC O8Y713; O8KYA9;
DT 28-FEB-2003 (Rel. 41; Created)
DI 28-FEB-2003 (Rel. 41; Last sequence update)
DI 28-FEB-2003 (Rel. 41; Last annotation update)
DE Hypothetical protein Lmo1284.
GN Lmo1284.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=EGD;
RX MEDLINE=22035523; PubMed=12039883;
RA Lampidis R., Kostrewa D., Hof H.;
R: *Molecular characterization of the genes encoding DNA gyrase and
R: topoisomerase IV of Listeria monocytogenes.;
RL J. Antimicrob. Chemother. 49:917-924(2002).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Deboux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Carrido P.,
RA Gautier L., Gobel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkta G.,

```

```

RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstrek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlucter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RL "Comparative genomics of Listeria species."
RL Science 294:849-852(2001).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE UPF0078 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AF084044; AAM48493.1; .
DR EMBL: AL591978; CAC99362.1; .
DR PIR: AD1235; AD1235.
DR Listlist: LM01284; .
DR HAMAP: MF_01043; .; 1.
DR Pfam: PF02660; DUF205; 1.
DR TIGRFAMs: TIGR00023; TIGR00023; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 5 27 POTENTIAL.
FT TRANSMEM 85 107 POTENTIAL.
FT TRANSMEM 114 136 POTENTIAL.
FT TRANSMEM 156 178 POTENTIAL.
FT CONFLICT 27 29 IFY -> FST (IN REF. 1).
FT CONFLICT 45 45 N -> I (IN REF. 1).
FT CONFLICT 62 62 D -> V (IN REF. 1).
SQ SEQUENCE 198 AA; 21599 MW; 7807B5406DFC5CD1 CRC64;

Query Match 3.5%; Score 7; DB 1; Length 198;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AVATSAG 36
DB 109 AVATSAG 115

RESULT 7
YD23 LISIN STANDARD; PRT; 198 AA.
AC Q9ZC58; 28-FEB-2003 (Rel. 41, Created)
DI 28-FEB-2003 (Rel. 41, Last sequence update)
DI 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein: Lin1323.
GN LIN1323.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Franche P., Bloeker H., Brandt P., Chakraborty T.,
RA Baquero F., Chetoui F., Couve E., de Bateau A., Deboix P.,
RA Charbit A., Chetoui F., Couve E., de Bateau A., Deboix P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fshi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haul C., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapatk S.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstrek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlucter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RL "Comparative genomics of Listeria species."
RL Science 294:849-852(2001).

```

```

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE UPF0078 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AL596168; CAC96554.1; .
DR PIR: AB1598; AB1598.
DR Listlist: LING1323; .
DR HAMAP: MF_01043; .; 1.
DR InterPro: IPR003811; DUF205.
DR Pfam: PF02660; DUF205; 1.
DR TIGRFAMs: TIGR00023; TIGR00023; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 5 27 POTENTIAL.
FT TRANSMEM 85 107 POTENTIAL.
FT TRANSMEM 114 136 POTENTIAL.
FT TRANSMEM 156 178 POTENTIAL.
SQ SEQUENCE 198 AA; 21632 MW; B161D1055R203406 CRC64;

Query Match 3.5%; Score 7; DB 1; Length 198;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AVATSAG 36
DB 109 AVATSAG 115

RESULT 8
YA35 STAEPP STANDARD; PRT; 202 AA.
AC P59253; 28-FEB-2003 (Rel. 41, Created)
DI 28-FEB-2003 (Rel. 41, Last sequence update)
DI 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein SE1035.
GN SE1035.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,
RA Chen Z., Wen Y.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE UPF0078 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AE016747; AA004632.1; .
DR HAMAP: MF_01043; .; 1.
DR Pfam: PF02660; DUF205; 1.
DR TIGRFAMs: TIGR00023; TIGR00023; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 4 26 POTENTIAL.
FT TRANSMEM 52 74 POTENTIAL.
FT TRANSMEM 84 106 POTENTIAL.
FT TRANSMEM 118 140 POTENTIAL.
FT TRANSMEM 141 161 POTENTIAL.

```

```

FI  TRANSMEM  118  140  POTENTIAL.
F1  TRANSMEM  141  161  POTENTIAL.
F1  TRANSMEM  162  181  POTENTIAL.
SQ  SEQUENCE  202 AA; 22232 MW; A9DA126B5731749C CRC64;

Query Match 3.5%; Score 7; DB 1; Length 202;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AVATSAG 36
    |||||
DQ 113 AVATSAG 119

RESULT 10
YC11_STRMU
ID YC11_STRMU STANDARD PRT: 212 AA.
AC P59255;
DQ 28-FEB-2003 (Rel. 41. Created)
LT 28-FEB-2003 (Rel. 41. Last sequence update)
DT 28-FEB-2003 (Rel. 41. Last annotation update)
DE Hypothetical protein SMU.1211.
GN SMU.1211.
GS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najaf F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT *Genome sequence of Streptococcus mutans UA159, a cariogenic dental
pathogen.;
RT Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
RL -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE UPF0078 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announc/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE014958; AAN58897.1; -
CC HAMAP: MF01043; -; 1.
CC DR TIGR: PF02660; DUF205; -
CC DR PF02660; TIGR00023; TIGR00023; 1.
CC KW Hypothetical protein; Transmembrane; Complete proteome.
CC F1 TRANSMEM 4 23 POTENTIAL.
CC F1 TRANSMEM 69 91 POTENTIAL.
CC F1 TRANSMEM 111 133 POTENTIAL.
CC F1 TRANSMEM 140 162 POTENTIAL.
CC F1 TRANSMEM 166 183 POTENTIAL.
CC F1 SEQUENCE 212 AA; 23417 MW; D44DA5F4270C5F5B CRC64;

Query Match 3.5%; Score 7; DB 1; Length 212;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AVATSAG 36
    |||||
DQ 105 AVATSAG 111

RESULT 11
YC30_STRAS
ID YC30_STRAS STANDARD PRT: 212 AA.
AC P59254;

```

```

DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein gbs1230/SAG1155.
CN GBS1230 OR SAG1155.
OS Streptococcus agalactiae (serotype III), and
OC Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX NCBI_TaxID=26495, 216466;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN-NEM316 / Serotype III;
RX MEDLINE-22242503; PubMed-12354221;
RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Frangou L.,
RA Msadek T., Zouine M., Couve E., Laliou L., Poyart C., Troadec P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
RT invasive neonatal disease."
RL Mol. Microbiol. 45:1499-1513(2002).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN-2603 V/R / Serotype V;
RX MEDLINE-2222988; PubMed-12200547;
RA Tettelin H., Masignani V., Cieslewicz M.J., Eisen J.A., Peterson S.,
RA Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.A.,
RA Radune D., Fedorova N.B., Scanlan D., Khouri H., Willigan S.,
RA Carthy A., Cline R.T., Van Aken S.E., Gill J., Scariselli M., Mora M.,
RA Iacobini E.T., Brettoni C., Galili G., Mariani M., Veqni F., Malone D.,
RA Rinaldo D., Rappuoli R., Telford J.L., Kaspar D.L., Grandi G.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V Streptococcus agalactiae."
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE UPF0078 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL766849; CAD46889.1; -
DR EMBL: AE014244; AAN0037.1; -
DR SAG1List: gbs1230; -
DR TIGR: SAG1155; -
DR HAMAP: MF_01043; -; 1.
DR Pfam: PF02660; DUF205; 1.
DR TIGRFAMs: TIGR00023; TIGR00023.1.
DR TIGRFAMs: TIGR00023; Transmembrane; Complete proteome.
FT TRANSMEM 4 23 POTENTIAL.
FT TRANSMEM 69 91 POTENTIAL.
FT TRANSMEM 113 135 POTENTIAL.
FT TRANSMEM 142 164 POTENTIAL.
FT TRANSMEM 168 185 POTENTIAL.
SQ SEQUENCE 212 AA; 23403 MW; 2BFE4A06F04E1657 CRC64;

Query Match 3.5%; Score 7; Dh 1; Length 212;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative C; Mismatches C; Indels 0; gaps C;

QY 30 AVATSAG 36
Db 105 AVATSAG 111
|||||||
RESULT 12
Y851_STRPN

```



```

FT TRANSMEM 2 22 POTENTIAL.
FT TRANSMEM 70 90 POTENTIAL.
FT TRANSMEM 112 132 POTENTIAL.
FT TRANSMEM 143 163 POTENTIAL.
FT TRANSMEM 164 184 POTENTIAL.
FT TRANSMEM 168 188 POTENTIAL.
FT CONFLICT 168 188 S -> F (IN REF. 1).
SQ SEQUENCE 213 AA: 22928 MW: 22CB089C17750818 CRC64;

Query Match 3.5% Score 7; DB 1: Length 213;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AVATSAG 36
DB 105 AVATSAG 111
|||||

RESULT 13
Y908_STRPY STANDARD: PRT: 213 AA.
ID Y908_STRPY STANDARD: PRT: 213 AA.
AC Q9A070; Q8K708;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein SPY0908/SPYM3_0623/SPS1230/spyM8_0566.
GN SPY0908 OR SPYM3_0623 OR SPS1230 OR SPYM8_0566.
OS Streptococcus pyogenes,
OS Streptococcus pyogenes (serotype M3), and
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314, 198466, 186103;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN-SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savić K., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White C.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.,
RT *Complete genome sequence of an M1 strain of Streptococcus pyogenes.*;
RT Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
RN [2];
RP SEQUENCE FROM N.A.
RC STRAIN-MGAS3115 / Serotype M3;
RX MEDLINE=21133808; PubMed=12122206;
RA Beres S.B., Silva G., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins J.F.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT *Genome sequence of a serotype M3 strain of group A Streptococcus;
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence.*;
RT Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
RN [3];
RP SEQUENCE FROM N.A.
RC STRAIN-SST-1 / Serotype M3;
RX Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,
RA Yamazaki K., Okahashi N., Kawabata S., Yasunaga T., Hattori M.,
RA Hayashi H., Hamada S.;
RT *The genome of invasive Streptococcus pyogenes; a comparative analysis
RT of S. pyogenes SST-1, SF370 and MGAS232.*;
RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [4];
RP SEQUENCE FROM N.A.
RC STRAIN-MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins J.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT *Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever

```

```

outbreaks.*;
RT Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE UPF0078 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE006539; AAK33824.1;
CC EMBL: AE014149; AAM79230.1; ALT_INIT.
CC EMBL: AF005144; BAC64325.1;
CC EMBL: AE010024; AAL97607.1;
CC HAMAP: MF_01043; -; 1.
CC InterPro: IPR003811; DUF205.
CC Pfam: PF02660; DUF205; 1.
CC TIGRfams: TIGR00023; TIGR00023; 1.
CC TRANSMEM 4 23 POTENTIAL.
CC TRANSMEM 68 90 POTENTIAL.
CC TRANSMEM 110 132 POTENTIAL.
CC TRANSMEM 139 161 POTENTIAL.
CC TRANSMEM 166 183 POTENTIAL.
CC SEQUENCE 213 AA: 23369 MW: 6A9881232A09766A CRC64;

Query Match 3.5% Score 7; DB 1: Length 213;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AVATSAG 36
DB 105 AVATSAG 111
|||||

RESULT 14
YKAC_LACLA STANDARD: PRT: 213 AA.
ID YKAC_LACLA STANDARD: PRT: 213 AA.
AC Q5CGW4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein ykac.
GN YKAC OR JL0978.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Mauget S., Jaillon O., Mialme K.,
RA Weissbach J., Ehrlich S.D., Sorokin A.;
RT *The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403.*;
RL Genome Res. 11:731-753(2001).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE UPF0078 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE006332; AAK05076.1;
CC HAMAP: MF_01043; -; 1.

```

```

DR InterPro: IPR003811; DUF205.
DR Pfam: PF02660; DUF205; 1.
DR TIGRFAMs: TIGR000023; TIGR000023; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 2 24 POTENTIAL.
FT TRANSMEM 44 63 POTENTIAL.
FT TRANSMEM 70 92 POTENTIAL.
FT TRANSMEM 112 134 POTENTIAL.
FT TRANSMEM 141 163 POTENTIAL.
FT TRANSMEM 167 184 POTENTIAL.
SQ SEQUENCE 213 AA; 23362 MW; 9672587000F31AE3 CRC64;

Query Match 3.5%; Score 7; DB 1; Length 215;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AVATSAG 36
DB 106 AVATSAG 112

RESULT 15
YPJC_BACSC STANDARD; PRT; 215 AA.
AC P42978;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DI 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein ypjc.
GN ypjc OR JOJC.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / Marburg;
RX MEDLINE=96349105; PubMed=8760912;
RA Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,
RA Serror P.;
RT *Sequence analysis of the Bacillus subtilis chromosome region between:
RL the strA and kgd loci cloned in a yeast artificial chromosome.*;
RN Microbiology 142:2005-2016(1996).
RP [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertsch M.C., Bessieres P., Bolotin A., Borchert S.,
RA Borris R., Boursier L., Braus A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connor I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Erington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi S.,
RA Guiseppi G., Guy B.J., Haga K., Haeck J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klager-Bianchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Potwolik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Serror S.J., Serror P., Shin B.S., Soldo E.,
RA Sorokin A., Tacconi E., Takagi I., Takahashi H., Takemaru K.,
RA Takeuchi M., Takakoshi A., Takata T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzengruber T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,

```

```

RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
subtilis.";
RL Nature 390:249-256(1997).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: TO B. SUBTILIS YITT AND YOFU.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L38424; AAA92872.1; -
DR EMBL: L47709; AAB38440.1; -
DR EMBL: Z99115; CAB14167.1; -
DR PIR: C69937; C69937.
DR Subtilist: BGL1209; ypjc.
DR InterPro: IPR003740; DUF151.
DR Pfam: PF02588; DUF161; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 31 51 POTENTIAL.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 104 124 POTENTIAL.
SQ SEQUENCE 215 AA; 23582 MW; D314CF7225F8A983 CRC64;

Query Match 3.5%; Score 7; DB 1; Length 215;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 VALSLFL 126
DB 10 VALSLFL 16

Search completed: September 26, 2003, 17:37:19
Job time : 26.0681 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2003, 17:29:43 : Search time 104.868 Seconds
(without alignments)
497.069 Million cell updates/sec

Title: us-09-852-100a-2_copy_68_269

Perfect score: 202

Sequence: 1 PSGPSAPEAVTARLVGVLMF.....TLRLSLTNETFRKTKQLYP 202

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 6

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

SPREMBL_23:*

1: sp.archaea:*

2: sp.bacteria:*

3: sp.fungi:*

4: sp.human:*

5: sp.invertebrate:*

6: sp.mammal:*

7: sp.mhc:*

8: sp.organella:*

9: sp.phage:*

10: sp.plant:*

11: sp.podent:*

12: sp.virus:*

13: sp.vertebrate:*

14: sp.unclassified:*

15: sp.virus:*

16: sp.bacteriap:*

17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	202	100.0	207	4 Q9BX74	Q9BX74 homo sapien
2	99	49.0	208	11 Q99MB3	Q99MB3 mus musculu
3	8	4.0	80	4 Q9H046	Q9H046 homo sapien
4	8	4.0	178	5 Q9W2H1	Q9W2H1 drosophila
5	8	4.0	195	5 Q95025	Q95025 caenorhabdi
6	8	4.0	206	2 Q93164	Q93164 rhodovulum
7	8	4.0	221	4 Q9H651	Q9H651 homo sapien
8	8	4.0	230	11 Q90156	Q90156 mus musculu
9	8	4.0	247	4 Q9B8N9	Q9B8N9 homo sapien
10	8	4.0	261	11 Q9B8H3	Q9B8H3 mus musculu
11	8	4.0	284	5 Q9U4H5	Q9U4H5 drosophila
12	8	4.0	548	5 Q19496	Q19496 caenorhabdi
13	8	4.0	1638	5 Q9VCW7	Q9VCW7 drosophila
14	7	3.5	38	11 Q8CAM1	Q8CAM1 mus musculu
15	7	3.5	97	16 Q8U561	Q8U561 agrobacteri
16	7	3.5	132	2 Q9S3F2	Q9S3F2 helicobacte

ALIGNMENTS

RESULT 1

Q9BX74 PRELIMINARY: PRT: 207 AA.
AC Q9BX74;
DT 01-JUN-2001 (TRENBLREL. 17, Created)
DT 01-JUN-2001 (TRENBLREL. 17, Last sequence update)
DE 01-OC-2002 (TRENBLREL. 22, Last annotation update)
DE Beta-amyloid binding protein.
GN BPP.
GS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
GX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278845;
RA Katkowski E.M., Lo C.F., Ning X., Walker S., Sofia H.J., Wang W.,
RA Edris W., Chanda P., Wagner E., Vile S., Ryan K., McHenry-Rinde R.,
RA Smith S.C., Wood A., Rhodes K.J., Kennedy J.D., Bard J.,
RA Jacobsen J.S., Ozenberger B.A.;
KT "Beta-Amyloid Peptide-Induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RL J. Biol. Chem. 276:16748-16756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE-Testis;
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF353990; AAK35064.1; -;
DR EMBL; BC029486; AAH29486.1; -;
KW Signal.
FT SIGNAL. 1 37 POTENTIAL.
SQ SEQUENCE 207 AA: 23226 MW: A5590FD7AECDF292 CRC64;

Query Match 100.0%; Score 202; DB 4; Length 207;

Best Local Similarity 100.0%; Pred. No. 4.3e-205;

Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSGPSAPEAVTARLVGVLMFVSVTTGPGAVATVSAGGESLKCEDLKVGQYICKDKIND 60


```

RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guerin H., Li P., Liao G., Miranda A., Mungall C.L.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Ceiniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AEO03453; AAF46720.1; -
DR EMBL: AY061343; AAL28891.1; -
DR FlyBase: FBgn034626; CG10795;
SQ SEQUENCE 178 AA; 19896 MW; 17C41166607ACC03 CRC64;

Query Match 4.0%; Score 8; DB 5; Length 178;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 DRYFLGYP 139
DB 108 DRYFLGYP 115
|||||||

RESULT 5
ID Q95Q25 PRELIMINARY; PRT; 195 AA.
AC Q95Q25;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 21.2 kDa protein.
GN C41D11.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorca; Rhabditida; Rhabditidae;
OC Rhabditidae; Peleoderinae; Caenorhabditis.
CX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=95069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018(1998);
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Gattung S., Magoi L.;
RT "The sequence of C. elegans cosmid C41D11."
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF003740; AAL08031.1; -
DR WormPeP: C41D11.9; CE29489;
KW Hypothetical protein.
SQ SEQUENCE 195 AA; 21203 MW; 35945F407F184DAE CRC64;

Query Match 4.0%; Score 8; DB 5; Length 195;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GADRFYLG 137
DB 148 GADRFYLG 155
|||||||

RESULT 6
Q93I64 PRELIMINARY; PRT; 206 AA.
ID Q93I64
AC Q93I64;

```

```

DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE BCHO (Fragment).
GN BCHO.
OS Rhodovulum sulfidophilum (Rhodobacter sulfidophilus).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacteriales;
OC Rhodobacteraceae; Rhodovulum.
CX NCBI_TaxID=35806;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W4;
RA Yoshida M., Masuda S., Nagashima K.V., Vermeiglio A., Shinada K.,
RA Matsuura K.;
RT "In vitro and in vivo electron transfer to the tritheme cytochrome
RT subunit bound to the photosynthetic reaction center complex in the
RT purple bacterium Rhodovulum sulfidophilum."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB050579; BAB70691.1; -
DR InterPro: IPR000073; A/B_hydrolase.
DR InterPro: IPR003089; AB_hydrolase.
DR InterPro: IPR000379; Ser_estrs_site.
DR Pfam: PF00561; abhydrolase; 1.
DR PRINTS: PR00111; ABHYDROLASE.
DR NON_TER 1;
FI NON_TER 1;
SQ SEQUENCE 206 AA; 21500 MW; 8F6713A03BB5B076 CRC64;

Query Match 4.0%; Score 8; DB 2; Length 206;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 APEAVTAR 13
DB 190 APEAVTAR 197
|||||||

RESULT 7
Q9H651 PRELIMINARY; PRT; 221 AA.
AC Q9H651;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ22604 (BBP-like protein 2).
GN BL22.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Uta T.,
RA Suzuki Y., Ohtsushi M., Nishi T., Shibahara T., Tanaka T.,
RA Nakamura Y., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC Kojowski E.M., Lo C.F., Ning X., Walker S., Sofia H.J., Wang W.,
RA Edris W., Chanda P., Wagner E., Vile S., Ryan K., McHenry-Rinde B.,
RA Smith S.C., Wood A., Rhodes K.J., Kennedy J.D., Bard J.,
RA Jacobsen J.S., Ozenberger B.A.;
RT "Beta-amyloid peptide-induced apoptosis regulated by a novel protein
RT containing a G protein activation module."
RL J. Biol. Chem. 0:0-0(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK026257; BAB15415.1; -
DR EMBL: AF353992; AAK35066.1; -

```

```

DR EMBL; BC008973; AA08873.1; -.
KW Hypothetical protein.
SQ SEQUENCE 221 AA; 24410 MW; 92151D6EF6363D74 CRC64;

Query Match
Best Local Similarity 4.0%; Score 8; DB 4; Length 221;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GADRFYLG 137
DB 174 GADRFYLG 181

RESULT 8
Q9D156 PRELIMINARY; PRT; 230 AA.
AC Q9D156
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DI 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 1110025109rik protein (RIKEN cDNA 1110025109 gene).
GN 1110025109RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh K., Ishii Y.,
RA Aizawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa K., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito K.,
RA Kadota K., Matsuda H.A., Asburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
RA Koenig P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirni L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli C., Mombacchi P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kotsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK003917; BAB23075.1; -.
DR EMBL; BC024620; AAH24620.1; -.
DR MGD; MGI:1915884; 1110025109rik.
SQ SEQUENCE 230 AA; 25639 MW; 396D65D8BEE99A5 CRC64;

Query Match
Best Local Similarity 4.0%; Score 8; DB 11; Length 230;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GADRFYLG 137
DB 183 GADRFYLG 190

RESULT 9
Q9BRN9 PRELIMINARY; PRT; 247 AA.
AC Q9BRN9
ID Q9BRN9
AC Q9BRN9;

```

```

DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Similar to hypothetical protein FLJ22604.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC006150; AAH06150.1; -.
KW Hypothetical protein.
SQ SEQUENCE 247 AA; 27161 MW; CE1D0D9C53DDF73C CRC64;

Query Match
Best Local Similarity 4.0%; Score 8; DB 4; Length 247;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GADRFYLG 137
DB 200 GADRFYLG 207

RESULT 10
CE5J83 PRELIMINARY; PRT; 261 AA.
AC CE5J83
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DI 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DI 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to Bsp-like protein 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Forelimb;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RT Nature 420:563-573(2002).
DR EMBL; AK077858; BAC37037.1; -.
DR EMBL; AK077858; BAC37037.1; -.
SQ SEQUENCE 261 AA; 28880 MW; 70346780D3CF5CDB CRC64;

Query Match
Best Local Similarity 4.0%; Score 8; DB 11; Length 261;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GADRFYLG 137
DB 214 GADRFYLG 22;

RESULT 11
Q9U4H5 PRELIMINARY; PRT; 294 AA.
AC Q9U4H5
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DI 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DI 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE BCNA.CH02974 (ALMONDEX) (AMX protein).
GN AMX OR BCNA.CH02974 OR CGI2127.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;

```

```

DQ      237 GADRFYLG 244

RESULT 12
OQ3496
ID      Q19496      PRELIMINARY;      PRT;      548 AA.
AC      Q19496;
DT      01-NOV-1996 (TrEMBLrel. 01, Created)
DT      01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT      01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE      Hypothetical 63.5 kDa protein.
GE      F16H11.5
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC      Rhabditidae; Pezodoriinae; Caenorhabditis.
OX      NCBI_TaxID=6239;
OX      (1)
RN      SEQUENCE FROM N.A.
RP      STRAIN-Bristol M2.
RP      MEDLINE=99069613; PubMed=9851916;
RX      None;
RA      "Genome sequence of the nematode C. elegans: a platform for
RA      investigating biology. The C. elegans Sequencing Consortium.";
RI      Science 282:2012-2018(1998).
RI      (2)
RP      SEQUENCE FROM N.A.
RP      STRAIN-Bristol M2;
RC      Wu X.;
RC      (3)
RC      "The sequence of C. elegans cosmid F16H11.";
RC      Submitted (APR-1996) to the EMBL/GenBank/DDJB databases.
RC      (1)
RC      SEQUENCE FROM N.A.
RC      STRAIN-Bristol M2;
RC      Waterston R.;
RA      "Direct Submission.";
RI      Submitted (AUG-2001) to the EMBL/GenBank/DDJB databases.
CC      1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC      1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
DR      EMBL; U55376; AAA98005.1; -.
DR      HSP; 220393; 1A6Y.
DR      WormPep; F16H11.5; C504400.
DR      InterPro; IPR000536; Hormone_rec_llg.
DR      InterPro; IPR001628; Znf_C4steroid.
DR      Pfam; PF00104; hormone_rec; 1.
DR      Pfam; PF00105; zf_C4; 1.
DR      PRINTS; PR00047; STROIDFINGER.
DR      ProDom; PD000035; Znf_C4steroid; 1.
DR      SMART; SM00430; HOLI; 1.
DR      SMART; SM00399; Znf_C4; 1.
KW      Hypothetical protein; DNA-binding; Metal-binding; Nuclear protein;
KW      Receptor; Transcription; Transcription regulation; Zinc; zinc-finger
SQ      SEQUENCE 548 AA; 63545 MW; 3BB5C51C6850BB96 CRC64;

Query Match      4.0%; Score 6; DB 5; Length 546;
Rest_local Similarity 100.0%; Pred No. 17;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps

OQ      66 KDSSGNET 93
ID      111111;
ID      109 KDSSGNET 116

RESULT 13
OQ9VCW7
ID      Q9VCW7      PRELIMINARY;      PRT;      1638 AA.
AC      Q9VCW7;
DT      01-MAY-2000 (TrEMBLrel. 13, Created)
DT      01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT      01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE      CG6954 protein.
GE      CG6954.
OS      Drosophila melanogaster (Fruit fly).
OC      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

```

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Koskins R.A., Galie R.E.,
 RA Sutton G.G., Richards S.R., Richards M.D., Zhang C., Chen L.X.,
 RA Brannon R.C., Rogers J.H.C., Blazer R.G., Chang M., Pfeiffer B.B.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Adayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 RA Ballwey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Danke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielson A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.C., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Itoqwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclob J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reider K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Stimpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao G.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.C.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003740; AAF56038.1; --
 DR FlyBase: FBgn0039000; CG6954.
 DR InterPro: IPR002110; ANK.
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR000159; RA_domain.
 DR Pfam: PF00023; ank; 2.
 DR Pfam: PF00041; fn3; 1.
 DR Pfam: PF00788; RA; 1.
 DR SMART: SM00248; ANK; 1.
 DR SMART: SM00060; FN3; 1.
 DR SMART: SM00314; FN3; 1.
 DR PROSITE: PS50088; ANK_REPEAT; 1.
 DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
 KW ANK repeat. Repeat.
 SQ SEQUENCE 1638 AA; 180548 MW; 3CC84F9A24B9F5DC CRC64;

Query Match 4.0%; Score 8; DB 5; Length 1638;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AVATSAGG 37

DB 1241 AVATSAGG 1248

|||||||

RESULT 14

Q8CAM1

ID Q8CAM1 PRELIMINARY: PRT: 38 AA.

Q8CAM1;
 CT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE LIM motif-containing protein kinase 2 (fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS7BL/6J; TISSUE=Hypothalamus;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RI Nature 420:563-573(2002).
 DR EMBL: AK038530; BAC30030.1; --
 FT NON_TER 1;
 SQ SEQUENCE 38 AA; 4188 MW; 2DB363A494415D42 CRC64;
 Query Match 3.5%; Score 7; DB 11; Length 38;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 ALSFLG 127
 DB 3 ALSFLG 9
 |||||

RESULT 15
 Q8U581 PRELIMINARY: PRT: 97 AA.
 IC Q8U581
 AC Q8U581;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DE AGR_C.3268p.
 DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 GN AGR_C.3268.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Grollio B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
 RA Houmlel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Dougherty D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens C58."
 RL Science 294:2323-2328(2001).
 DR EMBL: AE008099; AAK87547.1; --
 SQ SEQUENCE 97 AA; 10042 MW; 6B2B2E03D4137F36 CRC64;

Query Match 3.5%; Score 7; DB 16; Length 97;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 EAVTARL 14

DB 47 EAVTARL 53

|||||||

Search completed: September 26, 2003, 17:41:35
 Job time : 106.868 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2003, 17:22:38 : Search time 97 seconds
(without alignments)
715.630 Million cell updates/sec

Title: US-09-852-100A-2

Perfect score: 1439

Sequence: 1 MHLLKGNPNVIPAHCQKNT.....TRLTKLSITNEFRKIQ:YP 269

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp.archaea:*
- 2: sp.bacteria:*
- 3: sp.fungi:*
- 4: sp.human:*
- 5: sp.invertebrate:*
- 6: sp.mammal:*
- 7: sp.mhc:*
- 8: sp.organelle:*
- 9: sp.phage:*
- 10: sp.plant:*
- 11: sp.rodent:*
- 12: sp.virus:*
- 13: sp.vertebrate:*
- 14: sp.unclassified:*
- 15: sp.virus:*
- 16: sp.bacteriap:*
- 17: sp.bacteriap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1113	77.3	207	4	Q9BX74 Homo sapiens
2	941.5	65.4	208	11	Q99MB3 Q99MB3 mus musculus
3	338	23.5	178	5	Q9W2H1 Q9W2H1 drosophila
4	278	19.3	329	5	Q95FJ8 Q95FJ8 caenorhabditis
5	200	13.9	221	4	Q9H651 Q9H651 homo sapien
6	200	13.9	247	4	Q9BN9 Q9BN9 homo sapien
7	198.5	13.8	261	11	Q8BJ83 Q8BJ83 mus musculus
8	196	13.6	230	11	Q9D156 Q9D156 mus musculus
9	182	12.6	284	5	Q9U4H5 Q9U4H5 drosophila
10	172	12.0	80	4	Q9H046 Q9H046 homo sapien
11	169.5	11.8	195	5	Q95Q25 Q95Q25 caenorhabditis
12	136.5	9.5	214	4	Q9BX73 Q9BX73 homo sapien
13	125	8.7	172	5	Q9V786 Q9V786 drosophila
14	121	8.4	149	4	Q9BSR6 Q9BSR6 homo sapien
15	121	8.4	171	4	Q8N0X9 Q8N0X9 homo sapien
16	117	8.1	106	2	Q9S022 Q9S022 borrelia bu

Q9CWL9 mus musculus
Q8F0I4 mus musculus
Q8BJJ1 mus musculus
Q8F808 leptospira
Q8DJB9 synchococcus
Q9S015 borrelia bu
Q8F099 corynebacter
Q964R2 theileria t
Q932P1 staphylococ
Q75061 homo sapien
Q9RS03 deinococcus
Q8EVR7 mycoplasma
Q6668 equine herp
Q9V5W4 drosophila
Q961T2 drosophila
Q9J3E7 murine hepa
Q8K181 mus musculus
Q8K187 oryza sativ
Q22512 glycine max
Q98339 murine hepa
Q39227 murine hepa
Q9LLE7 streptomyce
Q9RZJ4 deinococcus
Q21535 caenorhabdi
Q9580 dunaliella
Q9J3F1 murine hepa
Q9VXJ5 drosophila
Q9KYT3 streptomyce
Q978F6 thermoplasm

ALIGNMENTS

RESULT 1

Q9BX74	PRELIMINARY;	PRT;	207 AA.
Q9BX74			
AC Q9BX74:			
DT 01-JUN-2001 (TrEMBLrel. 17, Created)			
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)			
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)			
DE Beta-amyloid binding protein.			
GN BAP.			
OS Homo sapiens (Human).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
CX NCBI_TaxID=9606;			
RN [1]			
RP MEDLINE=21276355; PubMed=11278849;			
RA Kajkowski E.M., Lo C.F., Ning X., Walker S., Sofia H.J., Wang W.,			
RA Edris W., Chanda P., Wagner E., Vile S., Ryan K., McHardy-Rinde H.,			
RA Smith S.C., Wood A., Rhodes K.J., Kennedy J.D., Bard J.,			
RA Jacobsen J.S., Ozenberger B.A.;			
RT "beta-Amyloid Peptide-induced Apoptosis Regulated by a Novel Protein			
RT Containing a G Protein Activation Module";			
RI J. Biol. Chem. 276:18748-18756(2001).			
RN [2]			
RP SEQUENCE FROM N.A.			
RA TISSUE=Testis;			
RA Strausberg R.;			
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.			
DR EMBL; AF353990; AAK35064.1; -			
DR EMBL; BC029486; AAK29486.1; -			
KW Signal.			
FT SIGNAL			
SC SEQUENCE 207 AA; 22326 MW; A5590FD7AECDF292 CRC64;			

POTENTIAL.

Query Match 77.3%; Score 1113; DB 4; Length 207;
Best Local Similarity 100.0%; Pred. No. 1.3e-94;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 MAAAHSPGSAPEAVTARLVGLVFWFVSITTPGWGAVA;SAGGEESLKCEDLVGVQICKD 122

88-76
PWA
185

```

Db      1  MAAAMPSPSAPEAVTARLVGLWLVSVTTGPGAVATISAGCESLKCED.KVGQYICKD 60
      123 PKINDATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFKP.SCRNVNGYS 182
      61 PKINDATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFKP.SCRNVNGYS 120
      183 YKVAVALSLFLGWLGDADRFYLGYPALGLLKFTCTVGFCGIGSLIDFILISMQIVGPSDGS 242
      121 YKVAVALSLFLGWLGDADRFYLGYPALGLLKFTCTVGFCGIGSLIDFILISMQIVGPSDGS 180
      243 YIIDYGGTRLRLSLTNETFRKTQLYP 269
      181 YIIDYGGTRLRLSLTNETFRKTQLYP 207

RESULT 2
Q99MB3 ID Q99MB3 PRELIMINARY: PRT: 208 AA.
AC Q99MB3;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Beta-amyloid binding protein.
GN BPP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=21276355; PubMed=11278849;
RA Kajkowski E.M., Lo C.F., Ning X., Walker S., Sofia H.J., Wang W.,
RA Edris W., Chanda P., Wagner E., Vile S., Ryan K., McHenry-Rinde H.,
RA Smith S.C., Wood A., Rhodes K.J., Kennedy J.D., Bard J.,
RA Jacobsen J.S., Ozenberger B.A.;
RT "beta-Amyloid Peptide-induced Apoptosis Regulated by a Novel Protein
RT Containing a G Protein Activation Module.";
RL J. Biol. Chem. 276:18748-18756(2001).
DR EMBL; AF353993; AAK35067.1; -;
DR MGD; MG1:2137022; Bbp.
SQ SEQUENCE 208 AA: 22271 MW: 91479321634F04C CRC64:

Query Match 65.4%; Score 941.5; DB 11: Length 208;
Best Local Similarity 85.1%; Pred. No. 8.4e-79;
Matches 177; Conservative 10; Mismatches 20; Indels 1; Gaps 1;

QY 63 MAAAMPSPSAPEAVTARLVGLWLVSVTTGPGAVATISA-GCESLKCEDLKVGQYICK 12;
Db 1 MAAAMPGRASPAAGPGGLRLTLVTVAAHGCAAGVGGEEPTKCEDLRVGQYICK 60
      122 DKINDATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFKP.SCRNVNGYS 181
      61 EPKINDATQEPVNCNTYTAHVSCFPAPNITCKDLSGNETHFTGSEVGLFKPISCRNVNGY 120
      182 SYKVAVALSLFLGWLGDADRFY-GYPALGLLKFTCTVGFCGIGSLIDFILISMQIVGPSDGS 241
      121 SYKVAVALSLFLGWLGDADRFYLGYPALGLLKFTCTVGFCGIGSLIDFILISMQIVGPSDGS 180
      242 YIIDYGGTRLRLSLTNETFRKTQLYP 269
      181 YIIDYGGTRLRLSLTNETFRKTQLYP 208

RESULT 3
Q9W2H1 ID Q9W2H1 PRELIMINARY: PRT: 178 AA.
AC Q9W2H1;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE CG10795 protein (LD27358P).

```

```

GN CG10795.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew K.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Doudin K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hootin D., Houston K.A., Howland T.J., Wei M.-H., Ibeqwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mallet B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nulton D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao C., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuccio J., Paclib J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003453; AAF46720.1; -;
DR EMBL; AY061343; AAL28891.1; -;
DR FlyBase; FBgn0034626; CG10795.
SQ SEQUENCE 178 AA: 19896 MW: 17C41166607ACC03 CRC64:

Query Match 23.5%; Score 338; DB 5: Length 178;
Best Local Similarity 42.6%; Pred. No. 3e-23;
Matches 69; Conservative 30; Mismatches 49; Indels 14; Gaps 5;

QY 107 SLKCEDLK-VGQYICKDP---KINDATQEPVNCNTY-TAHVSCFPAPNITCKDSSGNETH 161
      :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: |
DB 20 NVDCNELQMGGFCPPDPARGQIDPKTQOLAGCTREGRRYWCIAANEINCTF-TGNAT- 77
      162 FTGNEVGFKPISCRNVNGYSYKVAVALSLFLGWLGDADRFYLGYPALGLLKFTCTVGF 221
      78 -----FTREVPCKWTNGYHLDITLLLSVFLGWFVDRFYLGYPGLGLLAFCTLGGMFL 130

```



```

DL 90 PATNCSTSTKLLTKCSAHSVICHQGRNFYKRIKPCNWSSTSWIKTMI LSVVLGGGFA 149
QY 199 DREYLGYPALCGLLAFCTVCGFSGISGLDFILISMOIVCPGSSSVI 244
Db 150 DREYLGWLKSAIGLFLSGGLGVWTLVDVLIANGYIKPYDGSMTY 195

RESULT 12
Q9BX73 PRELIMINARY; PRT: 214 AA.
AC Q9BX73;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE BBP-like protein 1.
GN BLP1.
OS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA Rajkowsk E.M., Lo C.F., Ning X., Walker S., Sofia H.J., Wang W.,
RA Edris W., Chanda P., Wagner E., Vile S., Ryan K., McHenry-Rinde R.,
RA Smith S.C., Wood A., Rhodes K.J., Kennedy J.D., Bard J.,
RA Jacobsen J.S., Ozenberger B.A.;
RT "beta-Amyloid Peptide-induced Apoptosis Regulated by a Novel Protein
RT Containing a G Protein Activation Module.";
RL J. Biol. Chem. 276:19748-19756(2001).
DR EMBL; AF353991; AAK5065.1; -.
SQ SEQUENCE 214 AA; 22871 MW; BB328712AF2F78AB CRC64.

Query Match 9.5%; Score 136.5; Db 4; Length 214;
Best Local Similarity 27.8%; Pred. No. 0.00014;
Matches 58; Conservative 21; Mismatches 83; Indels 47; Gaps 10;

QY 59 SVSKMAAAMP-----SGPSAPEAVTARLVGVLFVSVTTPGKAVATSAGGES--LKCED 112
Db 33 SHSQNATAPELTISAGAOPE-----GPGASASWEYGDPSFVLASY 75

QY 113 LKVGQYICKDP--KINDAT--QE-PVNCNTNTAH-----VSCFFPAPNITCKDSSGN 156
Db 76 LPDEFIECDPDVHGVGNATASQELGYGCKFKFGGAYSDEHTSVQCHALDIEC---ASP 132

QY 159 ETHFTGNEVGFKPSICRNVCYSKVAVALSLFLGWLGVADREYLGYPALCGLLAFCTVGF 216
Db 133 RTFLREN-----KP--CKYTGHTFHTLLYSFFLGGCGVDRFCLGHTGTAVGKILLIGG 185

QY 219 CGIGSLIDFILISMOIVGPGSSSVIIDY 247
Db 186 LGIWMFVLDILLITGLGLMPDGSNNCTIVY 214

RESULT 13
Q9VY86 PRELIMINARY; PRT: 172 AA.
AC Q9VY86;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE CG11103 protein (LP03404p).
GN CG11103.
OS Drosophila melanogaster (Fruit fly).
CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Eteriyacta;
CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OX Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne C.D.,

```

```

Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brandon R.C., Rogers J.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
Abri J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Doup L.E., Dowies M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
Fosier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
Khalil M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
Mount S.M., Moy M., Murphy B., Murphy L., Muzzy D.M., Nelson D.L.,
Nelson D.R., Nelson X.A., Nixon K., Nusskern D.R., Paclab J.M.,
Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen R.,
Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Svirskas R., Tector C., Turner R., Venter F., Wang A.H., Wang X.,
Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong E.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.G.,
Gibbs S.A., Myers E.W., Rubin G.W., Venter J.C.;
X "The genome sequence of Drosophila melanogaster.";
X Science 287:2185-2195(2000).
X [2]
X SEQUENCE FROM N.A.
X Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
X Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
X Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
X Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
X Dodson K., Dorset V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
X Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
X Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
X Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
X McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
X Paclab J., Paragov S., Park S., Patel S., Pfeiffer B.,
X Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
X Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
X Williams S.M., Zaveri J.S., Smith H.G., Venter J.C., Rubin G.M.;
X "Sequencing of Drosophila melanogaster genome.";
X Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
X [3]
X SEQUENCE FROM N.A.
X Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
X Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
X Tu P.J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
X Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
X Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
X Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
X Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
X "Annotation of Drosophila melanogaster genome.";
X Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
X [4]
X SEQUENCE FROM N.A.
X Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
X Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
X [5]
X SEQUENCE FROM N.A.
X Flybase;
X Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

```

